





# **THE THEORY OF MATRICES**

BY

**F. R. GANTMACHER**

**VOLUME ONE**

**1959**

## PREFACE

THE MATRIX CALCULUS is widely applied nowadays in various branches of mathematics, mechanics, theoretical physics, theoretical electrical engineering, etc. However, neither in the Soviet nor the foreign literature is there a book that gives a sufficiently complete account of the problems of matrix theory and of its diverse applications. The present book is an attempt to fill this gap in the mathematical literature.

The book is based on lecture courses on the theory of matrices and its applications that the author has given several times in the course of the last seventeen years at the Universities of Moscow and Tiflis and at the Moscow Institute of Physical Technology.

The book is meant not only for mathematicians (undergraduates and research students) but also for specialists in allied fields (physics, engineering) who are interested in mathematics and its applications. Therefore the author has endeavoured to make his account of the material as accessible as possible, assuming only that the reader is acquainted with the theory of determinants and with the usual course of higher mathematics within the programme of higher technical education. Only a few isolated sections in the last chapters of the book require additional mathematical knowledge on the part of the reader. Moreover, the author has tried to keep the individual chapters as far as possible independent of each other. For example, Chapter V, *Functions of Matrices*, does not depend on the material contained in Chapters II and III. At those places of Chapter V where fundamental concepts introduced in Chapter IV are being used for the first time, the corresponding references are given. Thus, a reader who is acquainted with the rudiments of the theory of matrices can immediately begin with reading the chapters that interest him.

The book consists of two parts, containing fifteen chapters.

In Chapters I and III, information about matrices and linear operators is developed *ab initio* and the connection between operators and matrices is introduced.

Chapter II expounds the theoretical basis of Gauss's elimination method and certain associated effective methods of solving a system of  $n$  linear equations, for large  $n$ . In this chapter the reader also becomes acquainted with the technique of operating with matrices that are divided into rectangular 'blocks.'



In Chapter IV we introduce the extremely important 'characteristic' and 'minimal' polynomials of a square matrix, and the 'adjoint' and 'reduced adjoint' matrices.

In Chapter V, which is devoted to functions of matrices, we give the general definition of  $f(A)$  as well as concrete methods of computing it—where  $f(\lambda)$  is a function of a scalar argument  $\lambda$  and  $A$  is a square matrix. The concept of a function of a matrix is used in §§ 5 and 6 of this chapter for a complete investigation of the solutions of a system of linear differential equations of the first order with constant coefficients. Both the concept of a function of a matrix and this latter investigation of differential equations are based entirely on the concept of the minimal polynomial of a matrix and—in contrast to the usual exposition—do not use the so-called theory of elementary divisors, which is treated in Chapters VI and VII.

These five chapters constitute a first course on matrices and their applications. Very important problems in the theory of matrices arise in connection with the reduction of matrices to a normal form. This reduction is carried out on the basis of Weierstrass' theory of elementary divisors. In view of the importance of this theory we give two expositions in this book: an analytic one in Chapter VI and a geometric one in Chapter VII. We draw the reader's attention to §§ 7 and 8 of Chapter VI, where we study effective methods of finding a matrix that transforms a given matrix to normal form. In § 8 of Chapter VII we investigate in detail the method of A. N. Krylov for the practical computation of the coefficients of the characteristic polynomial.

In Chapter VIII certain types of matrix equations are solved. We also consider here the problem of determining all the matrices that are permutable with a given matrix and we study in detail the many-valued functions of matrices  $\sqrt[m]{A}$  and  $\ln A$ .

Chapters IX and X deal with the theory of linear operators in a unitary space and the theory of quadratic and hermitian forms. These chapters do not depend on Weierstrass' theory of elementary divisors and use, of the preceding material, only the basic information on matrices and linear operators contained in the first three chapters of the book. In § 9 of Chapter X we apply the theory of forms to the study of the principal oscillations of a system with  $n$  degrees of freedom. In § 11 of this chapter we give an account of Frobenius' deep results on the theory of Hankel forms. These results are used later, in Chapter XV, to study special cases of the Routh-Hurwitz problem.

The last five chapters form the second part of the book [the second volume, in the present English translation]. In Chapter XI we determine normal forms for complex symmetric, skew-symmetric, and orthogonal mat-

rices and establish interesting connections of these matrices with real matrices of the same classes and with unitary matrices.

In Chapter XII we expound the general theory of pencils of matrices of the form  $A + \lambda B$ , where  $A$  and  $B$  are arbitrary rectangular matrices of the same dimensions. Just as the study of regular pencils of matrices  $A + \lambda B$  is based on Weierstrass' theory of elementary divisors, so the study of singular pencils is built upon Kronecker's theory of minimal indices, which is, as it were, a further development of Weierstrass's theory. By means of Kronecker's theory—the author believes that he has succeeded in simplifying the exposition of this theory—we establish in Chapter XII canonical forms of the pencil of matrices  $A + \lambda B$  in the most general case. The results obtained there are applied to the study of systems of linear differential equations with constant coefficients.

In Chapter XIII we explain the remarkable spectral properties of matrices with non-negative elements and consider two important applications of matrices of this class: 1) homogeneous Markov chains in the theory of probability and 2) oscillatory properties of elastic vibrations in mechanics. The matrix method of studying homogeneous Markov chains was developed in the book [25] by V. I. Romanovskii and is based on the fact that the matrix of transition probabilities in a homogeneous Markov chain with a finite number of states is a matrix with non-negative elements of a special type (a 'stochastic' matrix).

The oscillatory properties of elastic vibrations are connected with another important class of non-negative matrices—the 'oscillation matrices.' These matrices and their applications were studied by M. G. Kreĭn jointly with the author of this book. In Chapter XIII, only certain basic results in this domain are presented. The reader can find a detailed account of the whole material in the monograph [7].

In Chapter XIV we compile the applications of the theory of matrices to systems of differential equations with variable coefficients. The central place (§§ 5-9) in this chapter belongs to the theory of the multiplicative integral (Produktintegral) and its connection with Volterra's infinitesimal calculus. These problems are almost entirely unknown in Soviet mathematical literature. In the first sections and in § 11, we study reducible systems (in the sense of Lyapunov) in connection with the problem of stability of motion; we also give certain results of N. P. Erugin. Sections 9-11 refer to the analytic theory of systems of differential equations. Here we clarify an inaccuracy in Birkhoff's fundamental theorem, which is usually applied to the investigation of the solution of a system of differential equations in the neighborhood of a singular point, and we establish a canonical form of the solution in the case of a regular singular point.

In § 12 of Chapter XIV we give a brief survey of some results of the fundamental investigations of I. A. Lappo-Danilevskii on analytic functions of several matrices and their applications to differential systems.

The last chapter, Chapter XV, deals with the applications of the theory of quadratic forms (in particular, of Hankel forms) to the Routh-Hurwitz problem of determining the number of roots of a polynomial in the right half-plane ( $\operatorname{Re} z > 0$ ). The first sections of the chapter contain the classical treatment of the problem. In § 5 we give the theorem of A. M. Lyapunov in which a stability criterion is set up which is equivalent to the Routh-Hurwitz criterion. Together with the stability criterion of Routh-Hurwitz we give, in § 11 of this chapter, the comparatively little known criterion of Liénard and Chipart in which the number of determinant inequalities is only about half of that in the Routh-Hurwitz criterion.

At the end of Chapter XV we exhibit the close connection between stability problems and two remarkable theorems of A. A. Markov and P. L. Chebyshev, which were obtained by these celebrated authors on the basis of the expansion of certain continued fractions of special types in series of decreasing powers of the argument. Here we give a matrix proof of these theorems.

This, then, is a brief summary of the contents of this book.

**F. R. Gantmacher**

## **PUBLISHERS' PREFACE**

THE PUBLISHERS wish to thank Professor Gantmacher for his kindness in communicating to the translator new versions of several paragraphs of the original Russian language book.

The Publishers also take pleasure in thanking the VEB Deutscher Verlag der Wissenschaften, whose many published translations of Russian scientific books into the German language include a counterpart of the present work, for their kind spirit of cooperation in agreeing to the use of their formulas in the preparation of the present work.

No material changes have been made in the text in translating the present work from the Russian except for the replacement of several paragraphs by the new versions supplied by Professor Gantmacher. Some changes in the references and in the Bibliography have been made for the benefit of the English-language reader.

## CONTENTS

PREFACE .....	iii
PUBLISHERS' PREFACE .....	vi
 <b>I. MATRICES AND OPERATIONS ON MATRICES</b> .....	 1
§ 1. Matrices. Basic notation .....	1
§ 2. Addition and multiplication of rectangular matrices.....	3
§ 3. Square matrices .....	12
§ 4. Compound matrices. Minors of the inverse matrix.....	19
 <b>II. THE ALGORITHM OF GAUSS AND SOME OF ITS APPLICATIONS</b> .....	 23
§ 1. Gauss's elimination method.....	23
§ 2. Mechanical interpretation of Gauss's algorithm.....	28
§ 3. Sylvester's determinant identity .....	31
§ 4. The decomposition of a square matrix into triangular factors .....	33
§ 5. The partition of a matrix into blocks. The technique of operating with partitioned matrices. The generalized algorithm of Gauss .....	41
 <b>III. LINEAR OPERATORS IN AN <math>n</math>-DIMENSIONAL VECTOR SPACE</b> .....	 50
§ 1. Vector spaces .....	50
§ 2. A linear operator mapping an $n$ -dimensional space into an $m$ -dimensional space .....	55
§ 3. Addition and multiplication of linear operators.....	57
§ 4. Transformation of coordinates .....	59
§ 5. Equivalent matrices. The rank of an operator. Sylvester's inequality .....	61
§ 6. Linear operators mapping an $n$ -dimensional space into itself .....	66

§ 7. Characteristic values and characteristic vectors of a linear operator .....	69
§ 8. Linear operators of simple structure.....	72
<b>IV. THE CHARACTERISTIC POLYNOMIAL AND THE MINIMAL POLYNOMIAL OF A MATRIX.....</b>	<b>76</b>
§ 1. Addition and multiplication of matrix polynomials.....	76
§ 2. Right and left division of matrix polynomials.....	77
§ 3. The generalized Bézout theorem.....	80
§ 4. The characteristic polynomial of a matrix. The adjoint matrix .....	82
§ 5. The method of Faddeev for the simultaneous computation of the coefficients of the characteristic polynomial and of the adjoint matrix .....	87
§ 6. The minimal polynomial of a matrix.....	89
<b>V. FUNCTIONS OF MATRICES.....</b>	<b>95</b>
§ 1. Definition of a function of a matrix.....	95
§ 2. The Lagrange-Sylvester interpolation polynomial.....	101
§ 3. Other forms of the definition of $f(A)$ . The components of the matrix $A$ .....	104
§ 4. Representation of functions of matrices by means of series	110
§ 5. Application of a function of a matrix to the integration of a system of linear differential equations with constant coefficients .....	116
§ 6. Stability of motion in the case of a linear system.....	125
<b>VI. EQUIVALENT TRANSFORMATIONS OF POLYNOMIAL MATRICES. ANALYTIC THEORY OF ELEMENTARY DIVISORS.....</b>	<b>130</b>
§ 1. Elementary transformations of a polynomial matrix.....	130
§ 2. Canonical form of a $\lambda$ -matrix.....	134
§ 3. Invariant polynomials and elementary divisors of a polynomial matrix .....	139
§ 4. Equivalence of linear binomials.....	145
§ 5. A criterion for similarity of matrices.....	147
§ 6. The normal forms of a matrix.....	149
§ 7. The elementary divisors of the matrix $f(A)$ .....	153



§ 8. A general method of constructing the transforming matrix	159
§ 9. Another method of constructing a transforming matrix.....	164

## **VII. THE STRUCTURE OF A LINEAR OPERATOR IN AN $n$ -DIMENSIONAL SPACE..... 175**

§ 1. The minimal polynomial of a vector and a space (with respect to a given linear operator).....	175
§ 2. Decomposition into invariant subspaces with co-prime minimal polynomials .....	177
§ 3. Congruence. Factor space .....	181
§ 4. Decomposition of a space into cyclic invariant subspaces.....	184
§ 5. The normal form of a matrix.....	190
§ 6. Invariant polynomials. Elementary divisors.....	193
§ 7. The Jordan normal form of a matrix.....	200
§ 8. Krylov's method of transforming the secular equation.....	202

## **VIII. MATRIX EQUATIONS..... 215**

§ 1. The equation $AX = XB$ .....	215
§ 2. The special case $A = B$ . Commuting matrices.....	220
§ 3. The equation $AX - XB = C$ .....	225
§ 4. The scalar equation $f(X) = 0$ .....	225
§ 5. Matrix polynomial equations .....	227
§ 6. The extraction of $m$ -th roots of a non-singular matrix.....	231
§ 7. The extraction of $m$ -th roots of a singular matrix.....	234
§ 8. The logarithm of a matrix.....	239

## **IX. LINEAR OPERATORS IN A UNITARY SPACE ..... 242**

§ 1. General considerations .....	242
§ 2. Metrization of a space.....	243
§ 3. Gram's criterion for linear dependence of vectors.....	246
§ 4. Orthogonal projection .....	248
§ 5. The geometrical meaning of the Gramian and some inequalities .....	250
§ 6. Orthogonalization of a sequence of vectors.....	256
§ 7. Orthonormal bases .....	262
§ 8. The adjoint operator.....	265

§ 9. Normal operators in a unitary space.....	268
§ 10. The spectra of normal, hermitian, and unitary operators.....	270
§ 11. Positive-semidefinite and positive-definite hermitian operators .....	274
§ 12. Polar decomposition of a linear operator in a unitary space. Cayley's formulas .....	276
§ 13. Linear operators in a euclidean space.....	280
§ 14. Polar decomposition of an operator and the Cayley formulas in a euclidean space.....	286
§ 15. Commuting normal operators .....	290
<b>X. QUADRATIC AND HERMITIAN FORMS.....</b>	<b>294</b>
§ 1. Transformation of the variables in a quadratic form.....	294
§ 2. Reduction of a quadratic form to a sum of squares. The law of inertia .....	296
§ 3. The methods of Lagrange and Jacobi of reducing a quadratic form to a sum of squares.....	299
§ 4. Positive quadratic forms .....	304
§ 5. Reduction of a quadratic form to principal axes.....	308
§ 6. Pencils of quadratic forms.....	310
§ 7. Extremal properties of the characteristic values of a regular pencil of forms .....	317
§ 8. Small oscillations of a system with $n$ degrees of freedom.....	326
§ 9. Hermitian forms .....	331
§ 10. Hankel forms .....	338
<b>BIBLIOGRAPHY .....</b>	<b>351</b>
<b>INDEX .....</b>	<b>369</b>

# CHAPTER I

## MATRICES AND OPERATIONS ON MATRICES

### § 1. Matrices. Basic Notation

1. Let  $F$  be a given number field.<sup>1</sup>

DEFINITION 1: A rectangular array of numbers of the field  $F$

$$\left\| \begin{array}{cccc} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{array} \right\| \quad (1)$$

is called a matrix. When  $m = n$ , the matrix is called square and the number  $m$ , equal to  $n$ , is called its order. In the general case the matrix is called rectangular (of dimension  $m \times n$ ). The numbers that constitute the matrix are called its elements.

NOTATION: In the double-subscript notation for the elements, the first subscript always denotes the row and the second subscript the column containing the given element.

As an alternative to the notation (1) for a matrix we shall also use the abbreviation

$$\|a_{ik}\| \quad (i = 1, 2, \dots, m; k = 1, 2, \dots, n). \quad (2)$$

Often the matrix (1) will also be denoted by a single letter, for example  $A$ . If  $A$  is a square matrix of order  $n$ , then we shall write  $A = \|a_{ik}\|_1^n$ . The determinant of a square matrix  $A = \|a_{ik}\|_1^n$  will be denoted by  $|a_{ik}|_1^n$  or by  $|A|$ .

---

<sup>1</sup> A number field is defined as an arbitrary collection of numbers within which the four operations of addition, subtraction, multiplication, and division by a non-zero number can always be carried out.

Examples of number fields are: the set of all rational numbers, the set of all real numbers, and the set of all complex numbers.

All the numbers that will occur in the sequel are assumed to belong to the number field given initially.

We introduce a concise notation for determinants formed from elements of the given matrix:

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \begin{vmatrix} a_{i_1 k_1} & a_{i_1 k_2} & \dots & a_{i_1 k_p} \\ a_{i_2 k_1} & a_{i_2 k_2} & \dots & a_{i_2 k_p} \\ \dots & \dots & \dots & \dots \\ a_{i_p k_1} & a_{i_p k_2} & \dots & a_{i_p k_p} \end{vmatrix}. \quad (3)$$

The determinant (3) is called a *minor* of  $A$  of order  $p$ , provided  $1 \leq i_1 < i_2 < \dots < i_p \leq m$  and  $1 \leq k_1 < k_2 < \dots < k_p \leq n$ . A rectangular matrix  $A = \| a_{ik} \|$  ( $i = 1, 2, \dots, m$ ;  $k = 1, 2, \dots, n$ ) has  $\binom{m}{p} \cdot \binom{n}{p}$  minors of order  $p$

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \quad \left( \begin{matrix} 1 \leq i_1 < i_2 < \dots < i_p \leq m \\ 1 \leq k_1 < k_2 < \dots < k_p \leq n \end{matrix} ; p \leq m, n \right). \quad (3')$$

The minors (3') in which  $i_1 = k_1, i_2 = k_2, \dots, i_p = k_p$  are called *principal minors*.

In the notation (3) the determinant of a square matrix  $A = \| a_{ik} \|_1^n$  can be written as follows:

$$|A| = A \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}.$$

The largest among the orders of the non zero minors generated by a matrix is called the *rank* of the matrix. If  $r$  is the rank of a rectangular matrix  $A$  of dimension  $m \times n$ , then obviously  $r \leq \min(m, n)$ .

A rectangular matrix consisting of a single column

$$\left\| \begin{matrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{matrix} \right\|$$

is called a *column matrix* and will be denoted by  $(x_1, x_2, \dots, x_n)$ .

A rectangular matrix consisting of a single row

$$\| z_1, z_2, \dots, z_n \|$$

is called a *row matrix* and will be denoted by  $[z_1, z_2, \dots, z_n]$ .

A square matrix in which all the elements outside the main diagonal are zero

$$\begin{vmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n \end{vmatrix}$$

is called a *diagonal matrix* and is denoted by<sup>2</sup>  $\|d_i \delta_{ik}\|_n$  or by

$$\{d_1, d_2, \dots, d_n\}.$$

Suppose that  $m$  quantities  $y_1, y_2, \dots, y_m$  have linear and homogeneous expressions in terms of  $n$  other quantities  $x_1, x_2, \dots, x_n$ :

$$\left. \begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ &\vdots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{aligned} \right\} \quad (4)$$

or more concisely,

$$y_i = \sum_{k=1}^n a_{ik} x_k \quad (i=1, 2, \dots, m). \quad (4')$$

The transformation of the quantities  $x_1, x_2, \dots, x_n$  into the quantities  $y_1, y_2, \dots, y_m$  by means of the formulas (4) is called a *linear transformation*.

The coefficients of this transformation form a rectangular matrix (1) of dimension  $m \times n$ .

The linear transformation (4) determines the matrix (1) uniquely, and vice versa.

In the next section we shall define the basic operations on rectangular matrices using the properties of the linear transformations (4) as our starting point.

## § 2. Addition and Multiplication of Rectangular Matrices

We shall define the basic operations on matrices: addition of matrices, multiplication of a matrix by a number, and multiplication of matrices.

**1.** Suppose that the quantities  $y_1, y_2, \dots, y_m$  are expressed in terms of the quantities  $x_1, x_2, \dots, x_n$  by means of the linear transformation

$$y_i = \sum_{k=1}^n a_{ik} x_k \quad (i=1, 2, \dots, m) \quad (5)$$

<sup>2</sup> Here  $\delta_{ik}$  is the Kronecker symbol:  $\delta_{ik} = \begin{cases} 1 & (i=k), \\ 0 & (i \neq k). \end{cases}$



and the quantities  $z_1, z_2, \dots, z_m$  in terms of the same quantities  $x_1, x_2, \dots, x_n$  by means of the transformation

$$z_i = \sum_{k=1}^n b_{ik} x_k \quad (i=1, 2, \dots, m). \quad (6)$$

Then

$$y_i + z_i = \sum_{k=1}^n (a_{ik} + b_{ik}) x_k \quad (i=1, 2, \dots, m). \quad (7)$$

In accordance with this, we formulate the following definition.

**DEFINITION 2:** The sum of two rectangular matrices  $A = \| a_{ik} \|$  and  $B = \| b_{ik} \|$ , both of dimension  $m \times n$ , is the matrix  $C = \| c_{ik} \|$ , of the same dimension, whose elements are the sums of the corresponding elements of the given matrices:

$$C = A + B,$$

where

$$c_{ik} = a_{ik} + b_{ik} \quad (i=1, 2, \dots, m; k=1, 2, \dots, n).$$

The operation of forming the sum of given matrices is called addition.

*Example.*

$$\left\| \begin{array}{ccc} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{array} \right\| + \left\| \begin{array}{ccc} c_1 & c_2 & c_3 \\ d_1 & d_2 & d_3 \end{array} \right\| = \left\| \begin{array}{ccc} a_1 + c_1 & a_2 + c_2 & a_3 + c_3 \\ b_1 + d_1 & b_2 + d_2 & b_3 + d_3 \end{array} \right\|.$$

According to Definition 2, only rectangular matrices of equal dimension can be added.

By virtue of the same definition, the coefficient matrix of the transformation (7) is the sum of the coefficient matrices of the transformations (5) and (6).

From the definition of matrix addition it follows immediately that this operation has the properties of commutativity and associativity:

1.  $A + B = B + A$ ;
2.  $(A + B) + C = A + (B + C)$ .

Here  $A$ ,  $B$ , and  $C$  are arbitrary rectangular matrices all of equal dimension.

The operation of addition of matrices extends in a natural way to the case of an arbitrary finite number of summands.

2. Let us multiply the quantities  $y_1, y_2, \dots, y_m$  in the transformation (5) by some number  $\alpha$  of  $\mathbb{F}$ . Then

$$\alpha y_i = \sum_{k=1}^n (\alpha a_{ik}) x_k \quad (i=1, 2, \dots, m).$$

In accordance with this, we formulate the following definition.

DEFINITION 3. The product of a matrix  $A = \| a_{ik} \|$  ( $i = 1, 2, \dots, m$ ;  $k = 1, 2, \dots, n$ ) by a number  $\alpha$  of  $\mathbb{F}$  is the matrix  $C = \| c_{ik} \|$  ( $i = 1, 2, \dots, m$ ;  $k = 1, 2, \dots, n$ ) whose elements are obtained from the corresponding elements of  $A$  by multiplication by  $\alpha$ :

$$C = \alpha A,$$

where

$$c_{ik} = \alpha a_{ik} \quad (i = 1, 2, \dots, m; k = 1, 2, \dots, n).$$

The operation of forming the product of a matrix by a number is called multiplication of the matrix by the number

Example.

$$\alpha \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} \alpha a_1 & \alpha a_2 & \alpha a_3 \\ \alpha b_1 & \alpha b_2 & \alpha b_3 \end{vmatrix}.$$

It is easy to see that

1.  $\alpha(A + B) = \alpha A + \alpha B$ ,
2.  $(\alpha + \beta)A = \alpha A + \beta A$ ,
3.  $(\alpha\beta)A = \alpha(\beta A)$ .

Here  $A$  and  $B$  are rectangular matrices of equal dimension and  $\alpha$  and  $\beta$  are numbers of  $\mathbb{F}$ .

The difference  $A - B$  of two rectangular matrices of equal dimension is defined by

$$A - B = A + (-1)B.$$

If  $A$  is a square matrix of order  $n$  and  $\alpha$  a number of  $\mathbb{F}$ , then<sup>3</sup>

$$|\alpha A| = \alpha^n |A|.$$

3. Suppose that the quantities  $z_1, z_2, \dots, z_m$  are expressed in terms of the quantities  $y_1, y_2, \dots, y_n$  by the transformation

$$z_i = \sum_{k=1}^n a_{ik} y_k \quad (i = 1, 2, \dots, m) \quad (8)$$

and that the quantities  $y_1, y_2, \dots, y_n$  are expressed in terms of the quantities  $x_1, x_2, \dots, x_q$  by the formulas

$$y_k = \sum_{j=1}^q b_{kj} x_j \quad (k = 1, 2, \dots, n) \quad (9)$$

Then on substituting these expressions for the  $y_k$  ( $k = 1, 2, \dots, n$ ) in (8) we can express  $z_1, z_2, \dots, z_m$  in terms of  $x_1, x_2, \dots, x_q$  by means of the composite transformation:

<sup>3</sup> Here the symbols  $|A|$  and  $|\alpha A|$  denote the determinants of the matrices  $A$  and  $\alpha A$  (see p. 1).

$$z_i = \sum_{k=1}^n a_{ik} \sum_{j=1}^q b_{kj} x_j = \sum_{j=1}^q \left( \sum_{k=1}^n a_{ik} b_{kj} \right) x_j \quad (i=1, 2, \dots, m). \quad (10)$$

In accordance with this we formulate the following definition.

DEFINITION 4. *The product of two rectangular matrices*

$$A = \begin{vmatrix} a_{11} & a_{12} \dots a_{1n} \\ a_{21} & a_{22} \dots a_{2n} \\ \dots & \dots \dots \dots \\ a_{m1} & a_{m2} \dots a_{mn} \end{vmatrix}, \quad B = \begin{vmatrix} b_{11} & b_{12} \dots b_{1q} \\ b_{21} & b_{22} \dots b_{2q} \\ \dots & \dots \dots \dots \\ b_{n1} & b_{n2} \dots b_{nq} \end{vmatrix}$$

is the matrix

$$C = \begin{vmatrix} c_{11} & c_{12} \dots c_{1q} \\ c_{21} & c_{22} \dots c_{2q} \\ \dots & \dots \dots \dots \\ c_{m1} & c_{m2} \dots c_{mq} \end{vmatrix}$$

in which the element  $c_{ij}$  at the intersection of the  $i$ -th row and the  $j$ -th column is the 'product'<sup>4</sup> of the  $i$ -th row of the first matrix  $A$  into the  $j$ -th column of the second matrix  $B$ :

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} \quad (i=1, 2, \dots, m; j=1, 2, \dots, q). \quad (11)$$

The operation of forming the product of given matrices is called *matrix multiplication*.

*Example.*

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \begin{vmatrix} c_1 & d_1 & e_1 & f_1 \\ c_2 & d_2 & e_2 & f_2 \\ c_3 & d_3 & e_3 & f_3 \end{vmatrix} = \\ = \begin{vmatrix} a_1 c_1 + a_2 c_2 + a_3 c_3 & a_1 d_1 + a_2 d_2 + a_3 d_3 & a_1 e_1 + a_2 e_2 + a_3 e_3 & a_1 f_1 + a_2 f_2 + a_3 f_3 \\ b_1 c_1 + b_2 c_2 + b_3 c_3 & b_1 d_1 + b_2 d_2 + b_3 d_3 & b_1 e_1 + b_2 e_2 + b_3 e_3 & b_1 f_1 + b_2 f_2 + b_3 f_3 \end{vmatrix}.$$

By Definition 4 the coefficient matrix of the transformation (10) is the product of the coefficient matrices of (8) and (9).

Note that the operation of multiplication of two rectangular matrices can only be carried out when the number of columns of the first factor is equal to the number of rows of the second. In particular, multiplication is always possible when both factors are square matrices of one and the same order.

<sup>4</sup> The product of two sequences of numbers  $a_1, a_2, \dots, a_n$  and  $b_1, b_2, \dots, b_n$  is defined as the sum of the products of the corresponding numbers:  $\sum_{i=1}^n a_i b_i$ .



Here  $x = (x_1, x_2, \dots, x_n)$  and  $y = (y_1, y_2, \dots, y_m)$  are column matrices and  $A = \|a_{ik}\|$  is a rectangular matrix of dimension  $m \times n$ .

Let us treat the special case when in the product  $C = AB$  the second factor is a square diagonal matrix  $B = \|d_1, d_2, \dots, d_n\|$ . Then it follows from (11) that

$$c_{ij} = a_{ij}d_j \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n),$$

i.e.,

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{vmatrix} \begin{vmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_n \end{vmatrix} = \begin{vmatrix} a_{11}d_1 & a_{12}d_2 & \dots & a_{1n}d_n \\ a_{21}d_1 & a_{22}d_2 & \dots & a_{2n}d_n \\ \dots & \dots & \dots & \dots \\ a_{m1}d_1 & a_{m2}d_2 & \dots & a_{mn}d_n \end{vmatrix}$$

Similarly,

$$\begin{vmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & d_m \end{vmatrix} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{vmatrix} = \begin{vmatrix} d_1a_{11} & d_1a_{12} & \dots & d_1a_{1n} \\ d_2a_{21} & d_2a_{22} & \dots & d_2a_{2n} \\ \dots & \dots & \dots & \dots \\ d_ma_{m1} & d_ma_{m2} & \dots & d_ma_{mn} \end{vmatrix}.$$

Hence: When a rectangular matrix  $A$  is multiplied on the right (left) by a diagonal matrix  $\{d_1, d_2, \dots\}$ , then the columns (rows) of  $A$  are multiplied by  $d_1, d_2, \dots$ , respectively.

4. Suppose that a square matrix  $C = \|c_{ij}\|_1^n$  is the product of two rectangular matrices  $A = \|a_{ik}\|$  and  $B = \|b_{kj}\|$  of dimension  $m \times n$  and  $n \times m$ , respectively:

$$\begin{vmatrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{vmatrix} = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{vmatrix} \begin{vmatrix} b_{11} & \dots & b_{1m} \\ b_{21} & \dots & b_{2m} \\ \dots & \dots & \dots \\ b_{n1} & \dots & b_{nm} \end{vmatrix}, \quad (12)$$

i.e.,

$$c_{ij} = \sum_{\alpha=1}^n a_{i\alpha}b_{\alpha j} \quad (i, j = 1, 2, \dots, m). \quad (13)$$



We shall establish the important *Binet-Cauchy formula*, which expresses determinant  $|C|$  in terms of the minors of  $A$  and  $B$ :

$$\begin{vmatrix} \dots & c_{1m} \\ \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{vmatrix} = \sum_{1 \leq k_1 < k_2 < \dots < k_m \leq n} \begin{vmatrix} a_{1k_1} & \dots & a_{1k_m} \\ \dots & \dots & \dots \\ a_{mk_1} & \dots & a_{mk_m} \end{vmatrix} \begin{vmatrix} b_{k_1 1} & \dots & b_{k_1 m} \\ \dots & \dots & \dots \\ b_{k_m 1} & \dots & b_{k_m m} \end{vmatrix} \quad (14)$$

or, in the notation<sup>5</sup> of page 2,

$$C \begin{pmatrix} 1 & 2 & \dots & m \\ 1 & 2 & \dots & m \end{pmatrix} = \sum_{1 \leq k_1 < k_2 < \dots < k_m \leq n} A \begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} B \begin{pmatrix} k_1 & k_2 & \dots & k_m \\ 1 & 2 & \dots & m \end{pmatrix}. \quad (14')$$

According to this formula the determinant of  $C$  is the sum of the products of all possible minors of the maximal ( $m$ -th) order<sup>5</sup> of  $A$  into the corresponding minors of the same order of  $B$ .

*Derivation of the Binet-Cauchy formula.* By (13) the determinant of  $C$  can be represented in the form

$$\begin{aligned} \begin{vmatrix} c_{11} & \dots & c_{1m} \\ \dots & \dots & \dots \\ c_{m1} & \dots & c_{mm} \end{vmatrix} &= \begin{vmatrix} \sum_{\alpha_1=1}^n a_{1\alpha_1} b_{\alpha_1 1} & \dots & \sum_{\alpha_m=1}^n a_{1\alpha_m} b_{\alpha_m m} \\ \dots & \dots & \dots \\ \sum_{\alpha_1=1}^n a_{m\alpha_1} b_{\alpha_1 1} & \dots & \sum_{\alpha_m=1}^n a_{m\alpha_m} b_{\alpha_m m} \end{vmatrix} \\ &= \sum_{\alpha_1, \dots, \alpha_m=1}^n \begin{vmatrix} a_{1\alpha_1} b_{\alpha_1 1} & \dots & a_{1\alpha_m} b_{\alpha_m m} \\ \dots & \dots & \dots \\ a_{m\alpha_1} b_{\alpha_1 1} & \dots & a_{m\alpha_m} b_{\alpha_m m} \end{vmatrix} \\ &= \sum_{\alpha_1, \dots, \alpha_m=1}^n A \begin{pmatrix} 1 & 2 & \dots & m \\ \alpha_1 & \alpha_2 & \dots & \alpha_m \end{pmatrix} b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_m m}. \end{aligned} \quad (15)$$

If  $m > n$ , then among the numbers  $\alpha_1, \alpha_2, \dots, \alpha_m$  there are always at least two that are equal, so that every summand on the right-hand side of (15) is zero. Hence in this case  $|C| = 0$ .

Now let  $m \leq n$ . Then in the sum on the right-hand side of (15) all those summands will be zero in which at least two of the subscripts  $\alpha_1, \alpha_2, \dots, \alpha_m$  are equal. All the remaining summands of (15) can be split into groups of  $m!$  terms each by combining into one group those summands that differ from each other only in the order of the subscripts  $\alpha_1, \alpha_2, \dots, \alpha_m$  (so that

<sup>5</sup> When  $m > n$ , the matrices  $A$  and  $B$  do not have minors of order  $m$ . In that case the right-hand sides of (14) and (14') are to be replaced by zero.

within each such group the subscripts  $\alpha_1, \alpha_2, \dots, \alpha_m$  have one and the same set of values). Now within one such group the sum of the corresponding terms is<sup>6</sup>

$$\begin{aligned} \sum \varepsilon(\alpha_1, \alpha_2, \dots, \alpha_m) A \begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_m m} = \\ = A \begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} \sum \varepsilon(\alpha_1, \alpha_2, \dots, \alpha_m) b_{\alpha_1 1} b_{\alpha_2 2} \dots b_{\alpha_m m} \\ = A \begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} B \begin{pmatrix} k_1 & k_2 & \dots & k_m \\ 1 & 2 & \dots & m \end{pmatrix}. \end{aligned}$$

Hence from (15) we obtain (14').

*Example 1.*

$$\left\| \begin{matrix} a_1 c_1 + a_2 c_2 + \dots + a_n c_n & a_1 d_1 + a_2 d_2 + \dots + a_n d_n \\ b_1 c_1 + b_2 c_2 + \dots + b_n c_n & b_1 d_1 + b_2 d_2 + \dots + b_n d_n \end{matrix} \right\| = \left\| \begin{matrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \end{matrix} \right\| \left\| \begin{matrix} c_1 & d_1 \\ c_2 & d_2 \\ \vdots & \vdots \\ c_n & d_n \end{matrix} \right\|.$$

Therefore formula (14) yields the so-called *Cauchy identity*

$$\left| \begin{matrix} a_1 c_1 + a_2 c_2 + \dots + a_n c_n & a_1 d_1 + a_2 d_2 + \dots + a_n d_n \\ b_1 c_1 + b_2 c_2 + \dots + b_n c_n & b_1 d_1 + b_2 d_2 + \dots + b_n d_n \end{matrix} \right| = \sum_{1 \leq i < k \leq n} \left| \begin{matrix} a_i & a_k \\ b_i & b_k \end{matrix} \right| \left| \begin{matrix} c_i & d_i \\ c_k & d_k \end{matrix} \right|. \quad (16)$$

Setting  $a_i = c_i, b_i = d_i$  ( $i = 1, 2, \dots, n$ ) in this identity, we obtain:

$$\left| \begin{matrix} a_1^2 + a_2^2 + \dots + a_n^2 & a_1 b_1 + a_2 b_2 + \dots + a_n b_n \\ a_1 b_1 + a_2 b_2 + \dots + a_n b_n & b_1^2 + b_2^2 + \dots + b_n^2 \end{matrix} \right| = \sum_{1 \leq i < k \leq n} \left| \begin{matrix} a_i & a_k \\ b_i & b_k \end{matrix} \right|^2.$$

If  $a_i$  and  $b_i$  ( $i = 1, 2, \dots, n$ ) are real numbers, we deduce the well-known inequality

$$(a_1 b_1 - a_2 b_2 + \dots + a_n b_n)^2 \leq (a_1^2 + a_2^2 + \dots + a_n^2) (b_1^2 + b_2^2 + \dots + b_n^2). \quad (17)$$

Here the equality sign holds if and only if all the numbers  $a_i$  are proportional to the corresponding numbers  $b_i$  ( $i = 1, 2, \dots, n$ ).

*Example 2.*

$$\left\| \begin{matrix} a_1 c_1 + b_1 d_1 & \dots & a_1 c_n + b_1 d_n \\ \vdots & \ddots & \vdots \\ a_n c_1 + b_n d_1 & \dots & a_n c_n + b_n d_n \end{matrix} \right\| = \left\| \begin{matrix} a_1 & b_1 \\ \vdots & \vdots \\ a_n & b_n \end{matrix} \right\| \left\| \begin{matrix} c_1 & \dots & c_n \\ d_1 & \dots & d_n \end{matrix} \right\|.$$

<sup>6</sup> Here  $k_1 < k_2 < \dots < k_m$  is the normal order of the subscripts  $\alpha_1, \alpha_2, \dots, \alpha_m$  and  $\varepsilon(\alpha_1, \alpha_2, \dots, \alpha_m) = (-1)^N$ , where  $N$  is the number of transpositions of the indices needed to put the permutation  $\alpha_1, \alpha_2, \dots, \alpha_m$  into normal order.

Therefore for  $n > 2$

$$\begin{vmatrix} a_1c_1 + b_1d_1 & \dots & a_1c_n + b_1d_n \\ \dots & \dots & \dots \\ a_nc_1 + b_nd_1 & \dots & a_nc_n + b_nd_n \end{vmatrix} = 0.$$

Let us consider the special case where  $A$  and  $B$  are square matrices of one and the same order  $n$ . When we set  $m = n$  in (14'), we arrive at the well-known multiplication theorem for determinants:

$$C \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} = A \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} B \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}$$

or, in another notation,

$$|C| = |AB| = |A| \cdot |B|. \quad (18)$$

Thus, the determinant of the product of two square matrices is equal to the product of the determinants of the factors.

5. The Binet-Cauchy formula enables us, in the general case also, to express the minors of the product of two rectangular matrices in terms of the minors of the factors. Let

$$A = \|a_{ik}\|, \quad B = \|b_{kj}\|, \quad C = \|c_{ij}\|$$

$$(i = 1, 2, \dots, m; \quad k = 1, 2, \dots, n; \quad j = 1, 2, \dots, q)$$

and

$$C = AB.$$

We consider an arbitrary minor of  $C$ :

$$C \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix} \quad \left( \begin{array}{l} 1 \leq i_1 < i_2 < \dots < i_p \leq m \\ 1 \leq j_1 < j_2 < \dots < j_p \leq q \end{array} ; \quad p \leq m \text{ and } p \leq q \right).$$

The matrix formed from the elements of this minor is the product of two rectangular matrices

$$\left\| \begin{array}{cccc} a_{i_1 1} & a_{i_1 2} & \dots & a_{i_1 n} \\ \dots & \dots & \dots & \dots \\ a_{i_p 1} & a_{i_p 2} & \dots & a_{i_p n} \end{array} \right\|, \quad \left\| \begin{array}{ccc} b_{1 j_1} & \dots & b_{1 j_p} \\ b_{2 j_1} & \dots & b_{2 j_p} \\ \dots & \dots & \dots \\ b_{n j_1} & \dots & b_{n j_p} \end{array} \right\|$$

Therefore, by applying the Binet-Cauchy formula, we obtain:

$$C \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix} = \sum_{1 \leq k_1 < k_2 < \dots < k_p \leq n} A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} B \begin{pmatrix} k_1 & k_2 & \dots & k_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix}. \quad (19)$$

For  $p = 1$  formula (19) goes over into (11). For  $p > 1$  formula (19) is a natural generalization of (11).

We mention another consequence of (19).

*The rank of the product of two rectangular matrices does not exceed the rank of either factor.*

If  $C = AB$  and  $r_A, r_B, r_C$  are the ranks of  $A, B, C$ , then

$$r_C \leq \min(r_A, r_B).$$

### § 3. Square Matrices

1. The square matrix of order  $n$  in which the main diagonal consists entirely of units and all the other elements are zero is called the *unit matrix* and is denoted by  $E^{(n)}$  or simply by  $E$ . The name 'unit matrix' is connected with the following property of  $E$ : For every rectangular matrix

$$A = \|a_{ik}\| \quad (i = 1, 2, \dots, m; k = 1, 2, \dots, n)$$

we have

$$E^{(m)}A = AE^{(n)} = A.$$

Clearly

$$E^{(n)} = \|\delta_{ik}\|_1^n.$$

Let  $A = \|a_{ik}\|_1^n$  be a square matrix. Then the *power of the matrix* is defined in the usual way:

$$A^p = \underbrace{AA \dots A}_{p \text{ times}} \quad (p = 1, 2, \dots); \quad A^0 = E.$$

From the associative property of matrix multiplication it follows that

$$A^p A^q = A^{p+q}$$

Here  $p$  and  $q$  are arbitrary non-negative integers.

---

<sup>7</sup> It follows from the Binet-Cauchy formula that the minors of order  $p$  in  $C$  for  $p > n$  (if minors of such orders exist) are all zero. In that case the right-hand side of (19) is to be replaced by zero. See footnote 5, p. 9.

We consider a polynomial (integral rational function) with coefficients in the field  $\mathbb{F}$ :

$$f(t) = \alpha_0 t^m + \alpha_1 t^{m-1} + \dots + \alpha_m.$$

Then by  $f(A)$  we shall mean the matrix

$$f(A) = \alpha_0 A^m + \alpha_1 A^{m-1} + \dots + \alpha_m E.$$

We define in this way a *polynomial in a matrix*.

Suppose that  $f(t)$  is the product of two polynomials  $g(t)$  and  $h(t)$ :

$$f(t) = g(t)h(t). \quad (21)$$

The polynomial  $f(t)$  is obtained from  $g(t)$  and  $h(t)$  by multiplication term by term and collection of similar terms. In this we make use of the multiplication rule for powers:  $t^p \cdot t^q = t^{p+q}$ . Since all these operations remain valid when the scalar  $t$  is replaced by the matrix  $A$ , it follows from (21) that

$$f(A) = g(A)h(A).$$

Hence, in particular,<sup>8</sup>

$$g(A)h(A) = h(A)g(A); \quad (22)$$

i.e., *two polynomials in one and the same matrix are always permutable*.

*Examples.*

Let the sequence of elements  $a_{ik}$  for which  $k - i = p$  ( $i - k = p$ ) in a rectangular matrix  $A = \|a_{ik}\|$  be called the  $p$ -th *superdiagonal* (*subdiagonal*) of the matrix. We denote by  $H^{(n)}$  the square matrix order  $n$  in which all the elements of the first superdiagonal are units and all the other elements are zero. The matrix  $H^{(n)}$  will also be denoted simply by  $H$ . Then

$$H = H^{(n)} = \begin{vmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{vmatrix}, \quad H^2 = \begin{vmatrix} 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & 1 \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & 0 & \dots & 0 \end{vmatrix}, \dots,$$

$$H^p = 0 \quad (p \geq n).$$

<sup>8</sup> Since each of these products is equal to one and the same  $f(A)$ , by virtue of the fact that  $h(t)g(t) = f(t)$ . It is worth mentioning that the substitution of matrices in an algebraic identity in several variables is not valid. The substitution of matrices that commute with one another, however, is allowable in this case.



By these equations, if

$$f(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_{n-1} t^{n-1} + \cdots$$

is a polynomial in  $t$ , then

$$f(H) = a_0 E + a_1 H + a_2 H^2 + \cdots = \begin{vmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ 0 & a_0 & a_1 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & a_2 \\ \vdots & \vdots & \vdots & \vdots & a_1 \\ 0 & 0 & 0 & \cdots & a_0 \end{vmatrix}$$

Similarly, if  $F$  is the square matrix of order  $n$  in which all the elements of the first subdiagonal are units and all others are zero, then

$$f(F) = a_0 E + a_1 F + a_2 F^2 + \cdots = \begin{vmatrix} a_0 & 0 & \cdots & 0 \\ a_1 & a_0 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & 0 \\ a_{n-1} & \cdots & a_1 & a_0 \end{vmatrix}.$$

We leave it to the reader to verify the following properties of the matrices  $H$  and  $F$ :

1. When an arbitrary rectangular matrix  $A$  of dimension  $m \times n$  is multiplied on the left by the matrix  $H$  (or  $F$ ) of order  $m$ , then all the rows of  $A$  are shifted upward (or downward) by one place, the first (last) row of  $A$  disappears, and the last (first) row of the product is filled by zeros. For example,

$$\begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix} \begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{vmatrix} = \begin{vmatrix} b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ 0 & 0 & 0 & 0 \end{vmatrix},$$

$$\begin{vmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} \begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 0 & 0 \\ a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \end{vmatrix}.$$

2. When an arbitrary rectangular matrix  $A$  of dimension  $m \times n$  is multiplied on the right by the matrix  $H$  (or  $F$ ) of order  $n$ , then all the columns of  $A$  are shifted to the right (left) by one place, the last (first) column of  $A$  disappears, and the first (last) column of the product is filled by zeros. For example,

$$\begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{vmatrix} \begin{vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{vmatrix} = \begin{vmatrix} 0 & a_1 & a_2 & a_3 \\ 0 & b_1 & b_2 & b_3 \\ 0 & c_1 & c_2 & c_3 \end{vmatrix},$$

$$\begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \end{vmatrix} \begin{vmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{vmatrix} = \begin{vmatrix} a_3 & a_2 & a_4 & 0 \\ b_2 & b_3 & b_4 & 0 \\ c_2 & c_3 & c_4 & 0 \end{vmatrix}.$$

**2.** A square matrix  $A$  is called *singular* if  $|A| = 0$ . Otherwise  $A$  is called *non-singular*.

Let  $A = \|a_{ik}\|_1^n$  be a non-singular matrix ( $|A| \neq 0$ ). Let us consider the linear transformation with coefficient matrix  $A$

$$y_i = \sum_{k=1}^n a_{ik} x_k \quad (i=1, 2, \dots, n). \quad (23)$$

When we regard (23) as equations for  $x_1, x_2, \dots, x_n$  and observe that the determinant of the system of equations (23) is, by assumption, different from zero, then we can express  $x_1, x_2, \dots, x_n$  in terms of  $y_1, y_2, \dots, y_n$  by means of the well-known formulas:

$$x_i = \frac{1}{|A|} \begin{vmatrix} a_{11} & \dots & a_{1,i-1} & y_1 & a_{1,i+1} & \dots & a_{1n} \\ a_{21} & \dots & a_{2,i-1} & y_2 & a_{2,i+1} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{n,i-1} & y_n & a_{n,i+1} & \dots & a_{nn} \end{vmatrix} = \sum_{k=1}^n a_{ik}^{(-1)} y_k \quad (i=1, 2, \dots, n). \quad (24)$$

We have thus obtained the 'inverse' transformation of the transformation (23). The coefficient matrix of this transformation

$$A^{-1} = \|a_{ik}^{(-1)}\|_1^n$$

will be called the *inverse matrix* of  $A$ . From (24) it is easy to see that

$$a_{ik}^{(-1)} = \frac{A_{ki}}{|A|} \quad (i, k=1, 2, \dots, n), \quad (25)$$

where  $A_{ki}$  is the algebraic complement (the cofactor) of the element  $a_{ki}$  in the determinant  $|A|$  ( $i, k=1, 2, \dots, n$ ).

For example, if

$$A = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} \text{ and } |A| \neq 0,$$

then

$$A^{-1} = \frac{1}{|A|} \begin{vmatrix} b_2c_3 - b_3c_2 & a_3c_2 - a_2c_3 & a_2b_3 - a_3b_2 \\ b_3c_1 - b_1c_3 & a_1c_3 - a_3c_1 & a_3b_1 - a_1b_3 \\ b_1c_2 - b_2c_1 & a_2c_1 - a_1c_2 & a_1b_2 - a_2b_1 \end{vmatrix}.$$

By forming the composite transformation of the given transformation (23) and the inverse (24), in either order, we obtain in both cases the identity transformation (with the unit matrix as coefficient matrix); therefore

$$AA^{-1} = A^{-1}A = E. \quad (26)$$

The validity of equation (26) can also be established by direct multiplication of the matrices  $A$  and  $A^{-1}$ . In fact, by (25) we have<sup>9</sup>

$$[AA^{-1}]_{ij} = \sum_{k=1}^n a_{ik}a_{kj}^{(-1)} = \frac{1}{|A|} \sum_{k=1}^n a_{ik}A_{jk} = \delta_{ij} \quad (i, j = 1, 2, \dots, n).$$

Similarly,

$$[A^{-1}A]_{ij} = \sum_{k=1}^n a_{ij}^{(-1)}a_{kj} = \frac{1}{|A|} \sum_{k=1}^n A_{ki}a_{kj} = \delta_{ij} \quad (i, j = 1, 2, \dots, n).$$

It is easy to see that the matrix equations

$$AX = E \text{ and } XA = E \quad (|A| \neq 0) \quad (27)$$

have no solutions other than  $X = A^{-1}$ . For by multiplying both sides of the first (second) equation on the left (right) by  $A^{-1}$  and using the associative property of matrix multiplication we obtain from (26) in both cases:<sup>10</sup>

$$X = A^{-1}.$$

<sup>9</sup> Here we make use of the well-known property of determinants that the sum of the products of the elements of an arbitrary column into the cofactors of the elements of that column is equal to the value of the determinant and the sum of the products of the elements of a column into the cofactors of the corresponding element of another column is zero.

<sup>10</sup> If  $A$  is a singular matrix, then the equations (27) have no solution. For if one of these equations had a solution  $X = \|x_{ik}\|_1^n$ , then we would have by the multiplication theorem of determinants (see formula (18)) that  $|A| \cdot |X| = |E| = 1$ , and this is impossible when  $|A| = 0$ .

In the same way it can be shown that each of the matrix equations

$$AX = B, XA = B \quad (|A| \neq 0), \quad (28)$$

where  $X$  and  $B$  are rectangular matrices of equal dimensions and  $A$  is a square matrix of appropriate order, have one and only one solution,

$$X = A^{-1}B \quad \text{and} \quad X = BA^{-1}, \quad (29)$$

respectively. The matrices (29) are the 'left' and the 'right' quotients on 'dividing'  $B$  by  $A$ . From (28) and (29) we deduce (see p. 12) that  $r_B \leq r_X$  and  $r_X \leq r_B$ , so that  $r_X = r_B$ . On comparing this with (28), we have:

*When a rectangular matrix is multiplied on the left or on the right by a non-singular matrix, the rank of the original matrix remains unchanged.*

Note that (26) implies  $|A| \cdot |A^{-1}| = 1$ , i.e.

$$|A^{-1}| = \frac{1}{|A|}.$$

For any two non-singular matrices we have

$$(AB)^{-1} = B^{-1}A^{-1}. \quad (30)$$

### 3. All the matrices of order $n$ form a ring<sup>11</sup> with unit element $E^{(n)}$ .

Since in this ring the operation of multiplication by a number of  $\mathfrak{F}$  is defined, and since there exists a basis of  $n^2$  linearly independent matrices in terms of which all the matrices of order  $n$  can be expressed linearly,<sup>12</sup> the ring of matrices of order  $n$  is an algebra.<sup>13</sup>

<sup>11</sup> A *ring* is a collection of elements in which two operations are defined and can always be carried out uniquely: the 'addition' of two elements (with the commutative and associative properties) and the 'multiplication' of two elements (with the associative and distributive properties with respect to addition); moreover, the addition is reversible. See, for example, van der Waerden, *Modern Algebra*, § 14.

<sup>12</sup> For, an arbitrary matrix  $A = \|a_{ik}\|_1^n$  with elements in  $\mathfrak{F}$  can be represented in the form  $A = \sum_{i,k=1}^n a_{ik} E_{ik}$ , where  $E_{ik}$  is the matrix of order  $n$  in which there is a 1 at the intersection of the  $i$ -th row and the  $k$ -th column and all the other elements are zeros.

<sup>13</sup> See, for example, van der Waerden, *Modern Algebra*, § 17.

All the square matrices of order  $n$  form a commutative group with respect to the operation of addition.<sup>14</sup> All the non-singular matrices of order  $n$  form a (non-commutative) group with respect to the operation of multiplication.

A square matrix  $A = \|a_{ik}\|_1^n$  is called *upper triangular* (*lower triangular*) if all the elements below (above) the main diagonal are zero:

$$A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn} \end{vmatrix}, \quad A = \begin{vmatrix} a_{11} & 0 & \dots & 0 \\ a_{21} & a_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

A diagonal matrix is a special case both of an upper triangular matrix and a lower triangular matrix.

Since the determinant of a triangular matrix is equal to the product of its diagonal elements, a triangular (and, in particular, a diagonal) matrix is non-singular if and only if all its diagonal elements are different from zero.

It is easy to verify that the sum and the product of two diagonal (upper triangular, lower triangular) matrices is a diagonal (upper triangular, lower triangular) matrix and that the inverse of a non-singular diagonal (upper triangular, lower triangular) matrix is a matrix of the same type. Therefore:

1. All the diagonal matrices of order  $n$  form a commutative group under the operation of addition, as do all the upper triangular matrices or all the lower triangular matrices.

2. All the non-singular diagonal matrices form a commutative group under multiplication.

3. All the non-singular upper (lower) triangular matrices form a (non-commutative) group under multiplication.

4. We conclude this section with a further important operation on matrices — *transposition*.

<sup>14</sup> A *group* is a set of objects in which an operation is defined which associates with any two elements  $a$  and  $b$  of the set a well-defined third element  $a * b$  of the same set provided that

- 1) the operation has the associative property  $((a * b) * c = a * (b * c))$ ,
- 2) there exists a unit element  $e$  in the set  $(a * e = e * a = a)$ , and
- 3) for every element  $a$  of the set there exists an inverse element  $a^{-1}$   $(a * a^{-1} = a^{-1} * a = e)$ .

A group is called *commutative*, or *abelian*, if the group operation has the commutative property. Concerning the group concept see, for example, [53], pp. 245ff.

If  $A = \| a_{ik} \|$  ( $i = 1, 2, \dots, m; k = 1, 2, \dots, n$ ), then the *transpose*  $A^T$  is defined as  $A^T = \| a_{ki}^T \|$ , where  $a_{ki}^T = a_{ik}$  ( $i = 1, 2, \dots, m; k = 1, 2, \dots, n$ ). If  $A$  is of dimension  $m \times n$ , then  $A^T$  is of dimension  $n \times m$ .

It is easy to verify the following properties:<sup>15</sup>

- a)  $(A + B)^T = A^T + B^T$ ,
- b)  $(\alpha A)^T = \alpha A^T$ ,
- c)  $(AB)^T = B^T A^T$ ,
- d)  $(A^{-1})^T = (A^T)^{-1}$ .

If a square matrix  $S = \| s_{ij} \|_1^n$  coincides with its transpose ( $S^T = S$ ), then it is called *symmetric*. In a symmetric matrix elements that are symmetrically placed with respect to the main diagonal are equal. Note that the product of two symmetric matrices is not, in general, symmetric. By 3., this holds if and only if the two given symmetric matrices are permutable.

If a square matrix  $K = \| k_{ij} \|_1^n$  differs from its transpose by a factor  $-1$  ( $K^T = -K$ ), then it is called *skew-symmetric*. In a skew-symmetric matrix any two elements that are symmetrical to the main diagonal differ from each other by a factor  $-1$  and the diagonal elements are zero. From 3. it follows that the product of two permutable skew-symmetric matrices is a symmetric matrix.<sup>16</sup>

#### § 4. Compound Matrices. Minors of the Inverse Matrix

1. Let  $A = \| a_{ik} \|_1^n$  be a given matrix. We consider all possible minors of  $A$  of order  $p$  ( $1 \leq p \leq n$ ):

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \left( 1 \leq i_1 < i_2 < \dots < i_p \leq n \right). \quad (31)$$

The number of these minors is  $N^2$ , where  $N = \binom{n}{p}$  is the number of combinations of  $n$  objects taken  $p$  at a time. In order to arrange the minors (31) in a square array, we enumerate in some definite order—lexicographic order, for example—all the  $N$  combinations of  $p$  indices selected from among the indices  $1, 2, \dots, n$ .

<sup>15</sup> In formulas 1., 2., 3.,  $A$  and  $B$  are arbitrary rectangular matrices for which the corresponding operations are feasible. In 4.,  $A$  is an arbitrary square non-singular matrix.

<sup>16</sup> As regards the representation of a square matrix  $A$  in the form of a product of two symmetric matrices ( $A = S_1 S_2$ ) or two skew-symmetric matrices ( $A = K_1 K_2$ ), see [357].

If the combinations of indices  $i_1 < i_2 < \dots < i_p$  and  $k_1 < k_2 < \dots < k_p$  have the numbers  $\alpha$  and  $\beta$ , then the minors (31) will also be denoted as follows:

$$a_{\alpha\beta} = A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix}.$$

By giving to  $\alpha$  and  $\beta$  independently all the values from 1 to  $N$ , we obtain all the minors of  $A = \|a_{ik}\|_1^n$  of order  $p$ .

The square matrix of order  $N$

$$\mathfrak{A}_p = \|a_{\alpha\beta}\|_1^N$$

is called the  $p$ -th *compound matrix* of  $A = \|a_{ik}\|_1^n$ ;  $p$  can take the values  $1, 2, \dots, n$ . Here  $\mathfrak{A}_1 = A$ , and  $\mathfrak{A}_n$  consists of the single element  $|A|$ .

*Note.* The order of enumeration of the combination of indices is fixed once and for all and does not depend on the choice of  $A$ .

*Example.* Let

$$A = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix}.$$

We enumerate all combinations of the indices 1, 2, 3, 4 taken two at a time by arranging them in the following order:

$$(12) \quad (13) \quad (14) \quad (23) \quad (24) \quad (34).$$

Then

$$\mathfrak{A}_2 = \begin{vmatrix} A \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} & A \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} & A \begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix} & A \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & A \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} & A \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \\ A \begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix} & A \begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix} & A \begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix} & A \begin{pmatrix} 1 & 3 \\ 2 & 3 \end{pmatrix} & A \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} & A \begin{pmatrix} 1 & 3 \\ 3 & 4 \end{pmatrix} \\ A \begin{pmatrix} 1 & 4 \\ 1 & 2 \end{pmatrix} & A \begin{pmatrix} 1 & 4 \\ 1 & 3 \end{pmatrix} & A \begin{pmatrix} 1 & 4 \\ 1 & 4 \end{pmatrix} & A \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} & A \begin{pmatrix} 1 & 4 \\ 2 & 4 \end{pmatrix} & A \begin{pmatrix} 1 & 4 \\ 3 & 4 \end{pmatrix} \\ A \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} & A \begin{pmatrix} 2 & 3 \\ 1 & 3 \end{pmatrix} & A \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} & A \begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix} & A \begin{pmatrix} 2 & 3 \\ 2 & 4 \end{pmatrix} & A \begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix} \\ A \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} & A \begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix} & A \begin{pmatrix} 2 & 4 \\ 1 & 4 \end{pmatrix} & A \begin{pmatrix} 2 & 4 \\ 2 & 3 \end{pmatrix} & A \begin{pmatrix} 2 & 4 \\ 2 & 4 \end{pmatrix} & A \begin{pmatrix} 2 & 4 \\ 3 & 4 \end{pmatrix} \\ A \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} & A \begin{pmatrix} 3 & 4 \\ 1 & 3 \end{pmatrix} & A \begin{pmatrix} 3 & 4 \\ 1 & 4 \end{pmatrix} & A \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix} & A \begin{pmatrix} 3 & 4 \\ 2 & 4 \end{pmatrix} & A \begin{pmatrix} 3 & 4 \\ 3 & 4 \end{pmatrix} \end{vmatrix}$$

We mention some properties of compound matrices:

1. *From  $C = AB$  it follows that  $\mathfrak{C}_p = \mathfrak{A}_p \cdot \mathfrak{B}_p$  ( $p = 1, 2, \dots, n$ ).*

For when we express the minors of order  $p$  ( $1 \leq p \leq n$ ) of the matrix product  $C$ , by formula (19), in terms of the minors of the same order of the factors, then we have:

$$C \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \sum_{1 \leq l_1 < l_2 < \dots < l_p \leq n} A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ l_1 & l_2 & \dots & l_p \end{pmatrix} B \begin{pmatrix} l_1 & l_2 & \dots & l_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \quad (1 \leq i_1 < i_2 < \dots < i_p \leq n) \quad (32)$$

Obviously, in the notation of this section, equation (32) can be written as follows:

$$c_{\alpha\beta} = \sum_{\lambda=1}^N a_{\alpha\lambda} b_{\lambda\beta} \quad (\alpha, \beta = 1, 2, \dots, N)$$

(here  $\alpha, \beta$ , and  $\lambda$  are the numbers of the combinations of indices  $i_1 < i_2 < \dots < i_p$ ;  $k_1 < k_2 < \dots < k_p$ ;  $l_1 < l_2 < \dots < l_p$ ). Hence

$$\mathfrak{C}_p = \mathfrak{A}_p \mathfrak{B}_p \quad (p = 1, 2, \dots, n).$$

2. *From  $B = A^{-1}$  it follows that  $\mathfrak{B}_p = \mathfrak{A}_p^{-1}$  ( $p = 1, 2, \dots, n$ ).*

This result follows immediately from the preceding one when we set  $C = E$  and bear in mind that  $\mathfrak{C}_p$  is the unit matrix of order  $N = \binom{n}{p}$

From 2. there follows an important formula that expresses the minors of the inverse matrix in terms of the minors of the given matrix:

If  $B = A^{-1}$ , then for arbitrary  $(1 \leq i_1 < i_2 < \dots < i_p \leq n)$

$$B \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \frac{(-1)^{\sum_{v=1}^p i_v + \sum_{v=1}^p k_v} A \begin{pmatrix} k'_1 & k'_2 & \dots & k'_{n-p} \\ i'_1 & i'_2 & \dots & i'_{n-p} \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}}, \quad (33)$$

where  $i_1 < i_2 < \dots < i_p$  and  $i'_1 < i'_2 < \dots < i'_{n-p}$  form a complete system of indices  $1, 2, \dots, n$ , as do  $k_1 < k_2 < \dots < k_p$  and  $k'_1 < k'_2 < \dots < k'_{n-p}$ .

For it follows from  $AB = E$  that

$$\mathfrak{A}_p \mathfrak{B}_p = \mathfrak{C}_p$$



or in more explicit form :

$$\sum_{\alpha=1}^N a_{\gamma\alpha} b_{\alpha\beta} = \delta_{\gamma\beta} = \begin{cases} 1 & (\gamma = \beta), \\ 0 & (\gamma \neq \beta). \end{cases} \quad (34)$$

Equations (34) can also be written as follows :

$$\sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} A \begin{pmatrix} j_1 & j_2 & \dots & j_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} B \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \begin{cases} 1, & \text{if } \sum_{v=1}^p (j_v - k_v)^2 = 0, \\ 0, & \text{if } \sum_{v=1}^p (j_v - k_v)^2 > 0 \end{cases} \quad (34')$$

$$\left( 1 \leq j_1 < j_2 < \dots < j_p \leq n \right).$$

On the other hand, when we apply the well-known Laplace expansion to the determinant  $|A|$ , we obtain

$$\sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} A \begin{pmatrix} j_1 & j_2 & \dots & j_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} \cdot (-1)^{\sum_{v=1}^p i_v + \sum_{v=1}^p k_v} A \begin{pmatrix} k'_1 & k'_2 & \dots & k'_{n-p} \\ i'_1 & i'_2 & \dots & i'_{n-p} \end{pmatrix} =$$

$$= \begin{cases} |A|, & \text{if } \sum_{v=1}^p (j_v - k_v)^2 = 0, \\ 0, & \text{if } \sum_{v=1}^p (j_v - k_v)^2 > 0, \end{cases} \quad (35)$$

where  $i_1 < i_2 < \dots < i_p$  and  $i'_1 < i'_2 < \dots < i'_{n-p}$  form a complete system of indices  $1, 2, \dots, n$ , as do  $k_1 < k_2 < \dots < k_p$  and  $k'_1 < k'_2 < \dots < k'_{n-p}$ . Comparison of (35) with (34') and (34) shows that the equations (34) are satisfied if we take together with  $b_{\alpha\beta}$  not  $B \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix}$  but rather

$$\frac{(-1)^{\sum_{v=1}^p i_v + \sum_{v=1}^p k_v} A \begin{pmatrix} k'_1 & k'_2 & \dots & k'_{n-p} \\ i'_1 & i'_2 & \dots & i'_{n-p} \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}}.$$

Since the elements  $b_{\alpha\beta}$  of the inverse matrix of  $\mathfrak{A}_p$  are uniquely determined by (34), equation (33) must hold.

## THE ALGORITHM OF GAUSS AND SOME OF ITS APPLICATIONS

1. Let

$$\left. \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= y_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= y_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= y_n \end{aligned} \right\} \quad (1)$$

In matrix form this system may be written as

$$Ax = y. \quad (1')$$

If  $A$  is non-singular, then we can rewrite this as

$$x = A^{-1}y, \quad (2)$$

or in explicit form :

$$x_i = \sum_{k=1}^n a_{ik}^{(-1)} y_k \quad (i=1, 2, \dots, n). \quad (2')$$

Thus, the task of computing the elements of the inverse matrix  $A^{-1} = \|\alpha_{ik}^{-1}\|_n^1$  is equivalent to the task of solving the system of equations (1) for arbitrary right-hand sides  $y_1, y_2, \dots, y_n$ . The elements of the inverse matrix are determined by the formulas (25) of Chapter I. However, the actual computation of the elements of  $A^{-1}$  by these formulas is very tedious for large  $n$ . Therefore, effective methods of computing the elements of an inverse matrix—and hence of solving a system of linear equations—are of great practical value.<sup>1</sup>

<sup>1</sup> For a detailed account of these methods, we refer the reader to the book by Faddeev [15] and the group of papers that appeared in *Uspehi Mat. Nauk*, Vol. 5, 3 (1950).

In the present chapter we expound the theoretical basis of some of these methods; they are variants of Gauss's elimination method, whose acquaintance the reader first made in his algebra course at school.

2. Suppose that in the system of equations (1) we have  $a_{11} \neq 0$ . We eliminate  $x_1$  from all the equations beginning with the second by adding to the second equation the first multiplied by  $-\frac{a_{21}}{a_{11}}$ , to the third the first multiplied by  $-\frac{a_{31}}{a_{11}}$ , and so on. The system (1) has now been replaced by the equivalent system

$$\left. \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= y_1 \\ a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n &= y_2^{(1)} \\ \dots &\dots \\ a_{r2}^{(r-1)}x_2 + \dots + a_{rn}^{(r-1)}x_n &= y_r^{(r-1)} \end{aligned} \right\} \quad (3)$$

The coefficients of the unknowns and the constant terms of the last  $n - 1$  equations are given by the formulas

$$a_{ij}^{(1)} = a_{ij} - \frac{a_{i1}}{a_{11}} a_{1j}, \quad y_i^{(1)} = y_i - \frac{a_{i1}}{a_{11}} y_1 \quad (i, j = 2, \dots, n). \quad (3')$$

Suppose that  $a_{22}^{(1)} \neq 0$ . Then we eliminate  $x_2$  in the same way from the last  $n - 2$  equations of the system (3) and obtain the system

$$\left. \begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n &= y_1 \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \dots + a_{2n}^{(1)}x_n &= y_2^{(1)} \\ a_{33}^{(2)}x_3 + \dots + a_{3n}^{(2)}x_n &= y_3^{(2)} \\ \dots &\dots \\ a_{rs}^{(r-1)}x_s + \dots + a_{rn}^{(r-1)}x_n &= y_r^{(r-1)} \end{aligned} \right\} \quad (4)$$

The new coefficients and the new right-hand sides are connected with the preceding ones by the formulas:

$$a_{ij}^{(2)} = a_{ij}^{(1)} - \frac{a_2^{(1)}}{a_{22}^{(1)}} a_{2j}^{(1)}, \quad y_i^{(2)} = y_i^{(1)} - \frac{a_2^{(1)}}{a_{22}^{(1)}} y_2^{(1)} \quad (i, j = 3, \dots, n). \quad (5)$$

Continuing the algorithm, we go in  $n - 1$  steps from the original system (1) to the triangular recurrent system

$$\left. \begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n &= y_1 \\ a_{22}^{(1)}x_2 + a_{23}^{(1)}x_3 + \dots + a_{2n}^{(1)}x_n &= y_2^{(1)} \\ a_{33}^{(2)}x_3 + \dots + a_{3n}^{(2)}x_n &= y_3^{(2)} \\ \vdots &\vdots \\ a_{nn}^{(n-1)}x_n &= y_n^{(n-1)}. \end{aligned} \right\} \quad (6)$$



From these formulas we find, by taking into account the structure (9) of  $G_p$ ,

$$A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} = a_{11} a_{22}^{(1)} \dots a_{pp}^{(p-1)}, \quad (11)$$

$$A \begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix} = a_{11} a_{22}^{(1)} \dots a_{pp}^{(p-1)} a_{ik}^{(p)} \quad (i, k = p+1, \dots, n), \quad (12)$$

When we divide the second of these equations by the first, we obtain the fundamental formulas<sup>2</sup>

$$a_{ik}^{(p)} = \frac{A \begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}} \quad (i, k = p+1, \dots, n). \quad (13)$$

If the conditions (7) hold for a given value of  $p$ , then they also hold for every smaller value of  $p$ . Therefore the formulas (13) are valid not only for the given value of  $p$  but also for all smaller values of  $p$ . The same holds true of (11). Hence instead of this formula we can write the equations

$$A \begin{pmatrix} 1 \\ 1 \end{pmatrix} = a_{11}, \quad A \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} = a_{11} a_{22}^{(1)}, \quad A \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = a_{11} a_{22}^{(1)} a_{33}^{(2)}, \dots \quad (14)$$

Thus, the conditions (7), i.e., the necessary and sufficient conditions for the feasibility of the first  $p$  steps in Gauss's algorithm, can be written in the form of the following inequalities:

$$A \begin{pmatrix} 1 \\ 1 \end{pmatrix} \neq 0, \quad A \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} \neq 0, \dots, \quad A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \neq 0 \quad (15)$$

From (14) we then find:

$$a_{11} = A \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad a_{22}^{(1)} = \frac{A \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}}{A \begin{pmatrix} 1 \\ 1 \end{pmatrix}}, \quad a_{33}^{(2)} = \frac{A \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}}{A \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}}, \dots, \quad a_{pp}^{(p-1)} = \frac{A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix}}. \quad (16)$$

In order to eliminate  $x_1, x_2, \dots, x_p$  consecutively by Gauss's algorithm it is necessary that all the values (16) should be different from zero, i.e., that the inequalities (15) should hold. However, the formulas for  $a_{ik}^{(p)}$  make sense if only the last of the conditions (15) holds.

<sup>2</sup> See [181], p. 89.

4. Suppose the coefficient matrix of the system of equations (1) to be of rank  $r$ . Then, by a suitable permutation of the equations and a renumbering of the unknowns, we can arrange that the following inequalities hold:

$$A \begin{pmatrix} 1 & 2 & \dots & j \\ 1 & 2 & \dots & j \end{pmatrix} \neq 0 \quad (j = 1, 2, \dots, r). \quad (17)$$

This enables us to eliminate  $x_1, x_2, \dots, x_r$  consecutively and to obtain the system of equations

$$\left. \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= y_1 \\ a_{22}^{(1)}x_2 + \dots + a_{2n}^{(1)}x_n &= y_2^{(1)} \\ &\dots \dots \dots \\ a_{rr}^{(r-1)}x_r + \dots + a_{rn}^{(r-1)}x_n &= y_r^{(r-1)} \\ a_{r+1,r+1}^{(r)}x_{r+1} + \dots + a_{r+1,n}^{(r)}x_n &= y_{r+1}^{(r)} \\ &\dots \dots \dots \\ a_{n,r+1}^{(r)}x_{r+1} + \dots + a_{nn}^{(r)}x_n &= y_n^{(r)} \end{aligned} \right\} \quad (18)$$

Here the coefficients are determined by the formulas (13). From these formulas it follows, because the rank of the matrix  $A = \|a_{ik}\|_1^n$  is equal to  $r$ , that

$$a_{ik}^{(r)} = 0 \quad (i, k = r+1, \dots, n). \quad (19)$$

Therefore the last  $n - r$  equations (18) reduce to the consistency conditions

$$y_i^{(r)} = 0 \quad (i = r+1, \dots, n). \quad (20)$$

Note that in the elimination algorithm the column of constant terms is subjected to the same transformations as the other columns, of coefficients. Therefore, by supplementing the matrix  $A = \|a_{ik}\|_1^n$  with an  $(n+1)$ -th column of the constant terms we obtain:

$$y_i^{(p)} = \frac{A \begin{pmatrix} 1 \dots p & i \\ 1 \dots p & n+1 \end{pmatrix}}{A \begin{pmatrix} 1 \dots p \\ 1 \dots p \end{pmatrix}} \quad (i = 1, 2, \dots, n; p = 1, 2, \dots, r). \quad (21)$$

In particular, the consistency conditions (20) reduce to the well-known equations

$$A \begin{pmatrix} 1 \dots r & r+j \\ 1 \dots r & n+1 \end{pmatrix} = 0 \quad (j = 1, 2, \dots, n-r). \quad (22)$$

If  $n = r$ , i.e. if the matrix  $A = \|a_{ik}\|_1^n$  is non-singular, and

$$A \begin{pmatrix} 1 & 2 & \dots & j \\ 1 & 2 & \dots & j \end{pmatrix} \neq 0 \quad (j = 1, 2, \dots, n),$$

then we can eliminate  $x_1, x_2, \dots, x_{n-1}$  in succession by means of Gauss's algorithm and reduce the system of equations to the form (6).

## § 2. Mechanical Interpretation of Gauss's Algorithm

1. We consider an arbitrary elastic statical system  $S$  supported on edges (for example, a string, a rod, a multispan rod, a membrane, a lamina, or a discrete system) and choose  $n$  points (1), (2),  $\dots$ , ( $n$ ) on it. We shall consider the displacements (sags)  $y_1, y_2, \dots, y_n$  of the points (1), (2),  $\dots$ , ( $n$ ) of  $S$  under the action of forces  $F_1, F_2, \dots, F_n$  applied at these points.

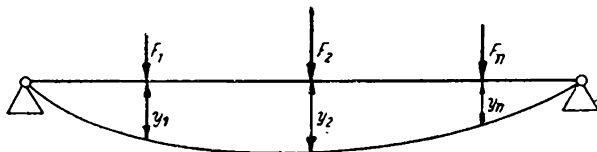


Fig. 1

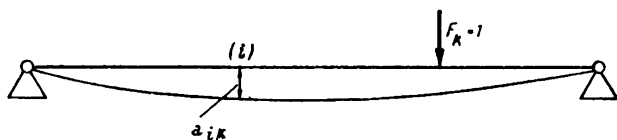


Fig. 2

We assume that the forces and the displacements are parallel to one and the same direction and are determined, therefore, by their algebraic magnitudes (Fig. 1). Moreover, we assume the *principle of linear superposition of forces*:

1. *Under the combined action of two systems of forces the corresponding displacements are added together.*

2. *When the magnitudes of all the forces are multiplied by one and the same real number, then all the displacements are multiplied by the same number.*





If

$$A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \neq 0,$$

then we can determine  $R_1, R_2, \dots, R_p$  from (25) and substitute the expressions so obtained in (24). This elimination of  $R_1, R_2, \dots, R_p$  can be carried out as follows. To the system of equations (25) we adjoin (24) written in the form

$$R_1 a_{i1} + \dots + R_p a_{ip} + a_{ik} - a_{ik}^{(p)} = 0. \quad (24')$$

Regarding (25) and (24') as a system of  $p+1$  homogeneous equations with non-zero solutions  $R_1, R_2, \dots, R_p, R_{p+1} = 1$ , we see that the determinant of the system must be zero:

$$\begin{vmatrix} a_{11} & \dots & a_{1p} & a_{1k} \\ \dots & \dots & \dots & \dots \\ a_{p1} & \dots & a_{pp} & a_{pk} \\ a_{i1} & \dots & a_{ip} & a_{ik} - a_{ik}^{(p)} \end{vmatrix} = 0.$$

Hence

$$a_{ik}^{(p)} = \frac{A \begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}} \quad (i, k = p+1, \dots, n). \quad (26)$$

These formulas express the coefficients of influence of the 'support' system  $S_p$  in terms of those of the original system  $S$ .

But formulas (26) coincide with formulas (13) of the preceding section. Therefore for every  $p (\leq n-1)$  the coefficients  $a_{ik}^{(p)}$  ( $i, k = p+1, \dots, n$ ) in the algorithm of Gauss are the coefficients of influence of the support system  $S_p$ .

The truth of this fundamental proposition can also be ascertained by purely mechanical considerations without recourse to the algebraic derivation of formulas (13). For this purpose we consider, to begin with, the special case of a single support:  $p=1$  (Fig. 3). In this case, the coefficients of influence of the system  $S_1$  are given by the formulas (we put  $p=1$  in (26)):

$$a_{ik}^{(1)} = \frac{A \begin{pmatrix} 1 & i \\ 1 & k \end{pmatrix}}{A \begin{pmatrix} 1 \\ 1 \end{pmatrix}} = a_{ik} - \frac{a_{i1}}{a_{11}} a_{1k} \quad (i, k = 1, 2, \dots, n)$$

These formulas coincide with the formulas (3').

Thus, if the coefficients  $a_{ik}$  ( $i, k = 1, 2, \dots, n$ ) in the system of equations (1) are the coefficients of influence of the statical system  $S$ , then the coefficients  $a_{ik}^{(1)}$  ( $i, k = 2, \dots, n$ ) in Gauss's algorithm are the coefficients of influence of the system  $S_1$ . Applying the same reasoning to the system  $S_1$  and introducing a second support at the point (2) in this system, we see that the coefficients  $a_{ik}^{(2)}$  ( $i, k = 3, \dots, n$ ) in the system of equations (4) are the coefficients of influence of the support system  $S_2$  and, in general, for every  $p$  ( $\leq n - 1$ ) the coefficients  $a_{ik}^{(p)}$  ( $i, k = p + 1, \dots, n$ ) in Gauss's algorithm are the coefficients of influence of the support system  $S_p$ .

From mechanical considerations it is clear that *the successive introduction of  $p$  supports is equivalent to the simultaneous introduction of these supports.*

*Note.* We wish to point out that in the mechanical interpretation of the elimination algorithm it was not necessary to assume that the points at which the displacements are investigated coincide with the points at which the forces  $F_1, F_2, \dots, F_n$  are applied. We can assume that  $y_1, y_2, \dots, y_n$  are the displacements of the points (1), (2),  $\dots$ , ( $n$ ) and that the forces  $F_1, F_2, \dots, F_n$  are applied at the points (1'), (2'),  $\dots$ , ( $n'$ ). Then  $a_{ik}$  is the coefficient of influence of the point ( $k'$ ) on the point ( $k$ ). In that case we must consider instead of the support at the point ( $j$ ) a generalized support at the points ( $j$ ), ( $j'$ ) under which the displacement at the point ( $j$ ) is maintained all the time equal to zero at the expense of a suitably chosen auxiliary force  $R_j$  at the point ( $j'$ ). The conditions that allow us to introduce  $p$  generalized supports at the points (1), (1'); (2), (2'),  $\dots$ ; ( $p$ ), ( $p'$ ), i.e., that allow us to satisfy the conditions  $y_1 = 0, y_2 = 0, \dots, y_p = 0$  for arbitrary  $F_{p+1}, \dots, F_n$  at the expense of suitable  $R_1 = F_1, \dots, R_p = F_p$ , can be expressed by the inequality

$$A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \neq 0.$$

### § 3. Sylvester's Determinant Identity

1. In § 1, a comparison of the matrices  $A$  and  $G_p$  led to equations (10) and (11).

These equations enable us to give an easy proof of the important *determinant identity of Sylvester*. For from (10) and (11) we find:

$$|A| = A \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} = A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \begin{vmatrix} a_{p+1, p+1}^{(p)} & \dots & a_{p+1, n}^{(p)} \\ \dots & \dots & \dots \\ a_{n, p+1}^{(p)} & \dots & a_{nn}^{(p)} \end{vmatrix}. \quad (27)$$

We introduce borderings of the minor  $A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}$  by the determinants

$$b_{ik} = A \begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix} \quad (i, k = p+1, \dots, n).$$

The matrix formed from these determinants will be denoted by

$$B = \|b_{ik}\|_{p+1}^n.$$

Then by formulas (13)

$$\begin{vmatrix} a_{p+1, p+1}^{(p)} & \dots & a_{p+1, n}^{(p)} \\ \vdots & \ddots & \vdots \\ a_{n, p+1}^{(p)} & \dots & a_{nn}^{(p)} \end{vmatrix} = \frac{\begin{vmatrix} b_{p+1, p+1} & \dots & b_{p+1, n} \\ \vdots & \ddots & \vdots \\ b_{n, p+1} & \dots & b_{nn} \end{vmatrix}}{\left[ A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \right]^{n-p}} = \frac{|B|}{\left[ A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \right]^{n-p}}.$$

Therefore equation (27) can be rewritten as follows:

$$|B| = \left[ A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \right]^{n-p-1} |A|. \quad (28)$$

This is Sylvester's determinant identity. It expresses the determinant  $|B|$  formed from the bordered determinants in terms of the original determinant and the bordered minor.

We have established equation (28) for a matrix  $A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}$  whose elements satisfy the inequalities

$$\begin{aligned} A \begin{pmatrix} 1 & 2 & \dots & j \\ 1 & 2 & \dots & j \end{pmatrix} &\neq 0 \\ (j = 1, 2, \dots, p). \end{aligned} \quad (29)$$

However, we can show by a 'continuity argument' that this restriction may be removed and that Sylvester's identity holds for an arbitrary matrix  $A = \|a_{ik}\|_1^n$ . For suppose that the inequalities (29) do not hold. We introduce the matrix

$$A_\varepsilon = A + \varepsilon E.$$

Obviously  $\lim_{\varepsilon \rightarrow 0} A_\varepsilon = A$ . On the other hand, the minors

$$\begin{aligned} A_\varepsilon \begin{pmatrix} 1 & 2 & \dots & j \\ 1 & 2 & \dots & j \end{pmatrix} &= \varepsilon^j + \dots \\ (j = 1, 2, \dots, p) \end{aligned}$$

are  $p$  polynomials in  $\varepsilon$  that do not vanish identically. Therefore we can choose a sequence  $\varepsilon_m \rightarrow 0$  such that

$$A_{\varepsilon_m} \begin{pmatrix} 1 & 2 & \dots & j \\ 1 & 2 & \dots & j \end{pmatrix} \neq 0 \quad (j = 1, 2, \dots, p; m = 1, 2, \dots).$$

We can write down the identity (28) for the matrices  $A_{\varepsilon_m}$ . Taking the limit  $m \rightarrow \infty$  on both sides of this identity, we obtain Sylvester's identity for the limit matrix<sup>3</sup>  $A = \lim_{m \rightarrow \infty} A_{\varepsilon_m}$ .

If we apply the identity (28) to the determinant

$$A \begin{pmatrix} 1 & 2 & \dots & p & i_1 & i_2 & \dots & i_q \\ 1 & 2 & \dots & p & k_1 & k_2 & \dots & k_q \end{pmatrix} \quad \left( p < i_1 < i_2 < \dots < i_q \leq n \right)$$

then we obtain a form of Sylvester's identity particularly convenient for applications

$$B \begin{pmatrix} i_1 & i_2 & \dots & i_q \\ k_1 & k_2 & \dots & k_q \end{pmatrix} = \left[ A \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \right]^{q-1} A \begin{pmatrix} 1 & 2 & \dots & p & i_1 & i_2 & \dots & i_q \\ 1 & 2 & \dots & p & k_1 & k_2 & \dots & k_q \end{pmatrix}. \quad (30)$$

#### § 4. The Decomposition of a Square Matrix into Triangular Factors

1. Let  $A = \|a_{ik}\|_1^n$  be a given matrix of rank  $r$ . We introduce the following notation for the successive principal minors of the matrix

$$D_k = A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \quad (k = 1, 2, \dots, n).$$

Let us assume that the conditions for the feasibility of Gauss's algorithm are satisfied:

$$D_k \neq 0 \quad (k = 1, 2, \dots, r).$$

We denote by  $G$  the coefficient matrix of the system of equations (18) to which the system

$$\sum_{k=1}^n a_{ik} x_k = y_i \quad (i = 1, 2, \dots, n)$$

<sup>3</sup> By the limit (for  $p \rightarrow \infty$ ) of a sequence of matrices  $X_p = \|x_{ik}^{(p)}\|_1^n$  we mean the matrix  $X = \|x_{ik}\|_1^n$ , where  $x_{ik} = \lim_{p \rightarrow \infty} x_{ik}^{(p)}$  ( $i, k = 1, 2, \dots, n$ ).

has been reduced by the elimination method of Gauss. The matrix  $G$  is of upper triangular form and the elements of its first  $r$  rows are determined by the formulas (13), while the elements of the last  $n - r$  rows are all equal to zero:<sup>4</sup>

$$G = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1r} & a_{1,r+1} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2r}^{(1)} & a_{2,r+1}^{(1)} & \dots & a_{2n}^{(1)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{rr}^{(r-1)} & a_{r,r+1}^{(r-1)} & \dots & a_{rn}^{(r-1)} \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{vmatrix}.$$

The transition from  $A$  to  $G$  is effected by a certain number  $N$  of operations of the following type: to the  $i$ -th row of the matrix we add the  $j$ -th row ( $j < i$ ), after a preliminary multiplication by some number  $\alpha$ . Such an operation is equivalent to the multiplication on the left of the matrix to be transformed by the matrix

$$\begin{matrix} & (j) & & (i) \\ \begin{vmatrix} 1 & \dots & 0 & \dots & 0 & \dots & 0 \\ \dots & & & & & & \\ \dots & & & & & & \\ \dots & & & & & & \\ \dots & & & 1 & & & \\ \dots & & & & & & \\ \dots & & & & & & \\ 0 & \dots & \alpha & \dots & 1 & \dots & 0 \\ \dots & & & & & & \\ \dots & & & & & & \\ 0 & \dots & 0 & \dots & 0 & \dots & 1 \end{vmatrix} \end{matrix} \quad (31)$$

In this matrix the main diagonal consists entirely of units, and all the remaining elements, except  $\alpha$ , are zero.

Thus,

$$G = W_N \dots W_2 W_1 A,$$

where each matrix  $W_1, W_2, \dots, W_N$  is of the form (31) and is therefore a lower triangular matrix with diagonal elements equal to 1.

<sup>4</sup> See formulas (19).  $G$  coincides with the matrix  $G_p$  (p. 25) for  $p = r$ .

Let.

$$W = W_N \dots W_2 W_1. \quad (32)$$

Then

$$G = WA. \quad (33)$$

We shall call  $W$  the *transforming matrix* for  $A$  in Gauss's elimination method. Both matrices  $G$  and  $W$  are uniquely determined by  $A$ . From (32) it follows that  $W$  is lower triangular with diagonal elements equal to 1.

Since  $W$  is non-singular, we obtain from (33):

$$A = W^{-1}G. \quad (33')$$

We have thus represented  $A$  in the form of a product of a lower triangular matrix  $W^{-1}$  and an upper triangular matrix  $G$ . The problem of decomposing a matrix  $A$  into factors of this type is completely answered by the following theorem:

**THEOREM 1:** Every matrix  $A = \|a_{ik}\|_1^n$  of rank  $r$  in which the first  $r$  successive principal minors are different from zero

$$D_k = A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \neq 0 \quad \text{for } k = 1, 2, \dots, r \quad (34)$$

can be represented in the form of a product of a lower triangular matrix  $B$  and an upper triangular matrix  $C$

$$A = BC = \begin{vmatrix} b_{11} & 0 & \dots & 0 \\ b_{21} & b_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{vmatrix} \begin{vmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ 0 & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & c_{nn} \end{vmatrix}. \quad (35)$$

Here

$$b_{11}c_{11} = D_1, \quad b_{22}c_{22} = \frac{D_2}{D_1}, \quad \dots, \quad b_{rr}c_{rr} = \frac{D_r}{D_{r-1}}. \quad (36)$$

The values of the first  $r$  diagonal elements of  $B$  and  $C$  can be chosen arbitrarily subject to the conditions (36).

When the first  $r$  diagonal elements of  $B$  and  $C$  are given, then the elements of the first  $r$  rows of  $B$  and of the first  $r$  columns of  $C$  are uniquely determined, and are given by the following formulas:

$$b_{gk} = b_{kk} \frac{A \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ 1 & 2 & \dots & k-1 & k \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}}, \quad c_{kg} = c_{kk} \frac{A \begin{pmatrix} 1 & 2 & \dots & k-1 & k \\ 1 & 2 & \dots & k-1 & g \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}} \quad (37)$$

$$(g = k, k+1, \dots, n; \quad k = 1, 2, \dots, r).$$

If  $r < n$  ( $|A| = 0$ ), then all the elements in the last  $n - r$  rows of  $B$  can be put equal to zero and all the elements of the last  $n - r$  columns of  $C$  can be chosen arbitrarily; or, conversely, the last  $n - r$  rows of  $C$  can be filled with zeros and the last  $n - r$  rows of  $B$  can be chosen arbitrarily.

*Proof.* That a representation of a matrix satisfying conditions (34) can be given in the form of a product (35) has been proved above (see (33')).

Now let  $B$  and  $C$  be arbitrary lower and upper triangular matrices whose product is  $A$ . Making use of the formulas for the minors of the product of two matrices we find:

$$A \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ 1 & 2 & \dots & k-1 & k \end{pmatrix} = \sum_{\alpha_1 < \alpha_2 < \dots < \alpha_k} B \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ \alpha_1 & \alpha_2 & \dots & \alpha_{k-1} & \alpha_k \end{pmatrix} C \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_k \\ 1 & 2 & \dots & k \end{pmatrix} \quad (38)$$

$$(g = k, k+1, \dots, n; k = 1, 2, \dots, r).$$

Since  $C$  is an upper triangular matrix, the first  $k$  columns of  $C$  contain only one non-vanishing minor of order  $k$ , namely  $C \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}$ . Therefore, equation (38) can be written as follows:

$$A \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ 1 & 2 & \dots & k-1 & k \end{pmatrix} = B \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ 1 & 2 & \dots & k-1 & k \end{pmatrix} C \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}$$

$$= b_{11} b_{22} \dots b_{k-1, k-1} b_{gk} c_{11} c_{22} \dots c_{kk} \quad (39)$$

$$(g = k, k+1, \dots, n; k = 1, 2, \dots, r).$$

We put  $g = k$  in this equation, obtaining

$$b_{11} b_{22} \dots b_{kk} c_{11} c_{22} \dots c_{kk} = D_k \quad (k = 1, 2, \dots, r), \quad (40)$$

and relations (36) follow.

Without violating equation (35) we may multiply the matrix  $B$  in that equation on the right by an arbitrary non-singular diagonal matrix  $M = \|\mu_i \delta_{ik}\|_1^n$ , while multiplying  $C$  at the same time on the left by  $M^{-1} = \|\mu_i^{-1} \delta_{ik}\|_1^n$ . But this is equivalent to multiplying the columns of  $B$  by  $\mu_1, \mu_2, \dots, \mu_n$ , respectively, and the rows of  $C$  by  $\mu_1^{-1}, \mu_2^{-1}, \dots, \mu_n^{-1}$ . We may therefore give arbitrary values to the diagonal elements  $b_{11}, b_{22}, \dots, b_{rr}$  and  $c_{11}, c_{22}, \dots, c_{rr}$ , provided they satisfy (36).

Further, from (39) and (40) we find:

$$b_{gk} = b_{kk} \frac{A \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ 1 & 2 & \dots & k-1 & k \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}} \quad (g = k, k+1, \dots, n; k = 1, 2, \dots, r),$$

i.e., the first formulas in (37). The second formulas in (37), for the elements of  $C$ , are established similarly.

We observe that in the multiplication of  $B$  and  $C$  the elements  $b_{kg}$  of the last  $n - r$  columns of  $B$  and the elements  $c_{gk}$  of the last  $n - r$  rows of  $C$  are multiplied only among each other. We have seen that all the elements of the last  $n - r$  rows of  $C$  may be chosen to be zero.<sup>5</sup> But as a consequence, the elements of the last  $n - r$  columns of  $B$  may be chosen arbitrarily. Clearly the product of  $B$  and  $C$  does not change if we choose the last  $n - r$  columns of  $B$  to be zeros and choose the elements of the last  $n - r$  rows of  $C$  arbitrarily.

This completes the proof of the theorem.

From this theorem there follow a number of interesting corollaries.

**COROLLARY 1:** *The elements of the first  $r$  columns of  $B$  and the first  $r$  rows of  $C$  are connected with the elements of  $A$  by the recurrence relations*

$$\left. \begin{aligned} b_{ik} &= \frac{a_{ik} - \sum_{j=1}^{k-1} b_{ij}c_{jk}}{c_{kk}} & (i \geq k; i = 1, 2, \dots, n; k = 1, 2, \dots, r), \\ c_{ik} &= \frac{a_{ik} - \sum_{j=1}^{i-1} b_{ij}c_{jk}}{b_{ii}} & (i \leq k; i = 1, 2, \dots, r; k = 1, 2, \dots, n). \end{aligned} \right\} \quad (41)$$

The relations (41) follow immediately from the matrix equation (35); they can be used to advantage in the actual computation of the elements of  $B$  and  $C$ .

**COROLLARY 2:** *If  $A = \|a_{ik}\|_1^n$  is a non-singular matrix ( $r = n$ ) satisfying (34), then the matrices  $B$  and  $C$  in the representation (35) are uniquely determined as soon as the diagonal elements of these matrices are chosen in accordance with (36).*

**COROLLARY 3:** *If  $S = \|s_{ik}\|_1^n$  is a symmetric matrix of rank  $r$  and*

$$D_k = S \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \neq 0 \quad (k = 1, 2, \dots, r),$$

then

$$S = B B',$$

where  $B = \|b_{ik}\|_1^n$  is a lower triangular matrix in which

<sup>5</sup> This follows from the representation (33'). Here, as we have shown already, arbitrary values may be given to the diagonal elements  $b_{11}, \dots, b_{rr}, c_{11}, \dots, c_{rr}$  provided (36) is satisfied by the introduction of suitable factors  $\mu_1, \mu_2, \dots, \mu_r$ .



$$b_{gk} = \begin{cases} \frac{1}{\gamma D_k D_{k-1}} A \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ 1 & 2 & \dots & k-1 & k \end{pmatrix} & (g = k, k+1, \dots, n; k = 1, 2, \dots, r), \\ 0 & (g = k, k+1, \dots, n; k = r+1, \dots, n). \end{cases} \quad (42)$$

2. In the representation (35) let the elements of the last  $n - r$  columns of  $C$  be zero. Then we may set

$$B = F \begin{vmatrix} b_{11} & & & 0 \\ & \ddots & & \\ & & b_{rr} & 0 \\ 0 & & & \ddots & \\ & & & & 0 \end{vmatrix}, \quad C = \begin{vmatrix} c_{11} & & & 0 \\ & \ddots & & \\ & & c_{rr} & 0 \\ 0 & & & \ddots & \\ & & & & 0 \end{vmatrix} L, \quad (43)$$

where  $F$  and  $L$  are upper and lower triangular matrices respectively; the first  $r$  diagonal elements of  $F$  and  $L$  are 1 and the elements of the last  $n - r$  columns of  $F$  and the last  $n - r$  rows of  $L$  can be chosen completely arbitrarily. Substituting (43) for  $B$  and  $C$  in (35) and using (36), we obtain the following theorem:

THEOREM 2: Every matrix  $A = \|a_{ik}\|_1^n$  of rank  $r$  in which

$$D_k = A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \neq 0 \quad \text{for } k = 1, 2, \dots, r$$

can be represented in the form of a product of a lower triangular matrix  $F$ , a diagonal matrix  $D$ , and an upper triangular matrix  $L$ :

$$A = FDL = \begin{vmatrix} 1 & 0 & \dots & 0 \\ f_{n1} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1} & f_{n2} & \dots & 1 \end{vmatrix} \begin{vmatrix} D_1 & & & \\ & D_2 & & \\ & & \ddots & \\ & & & D_r \\ & & & & 0 \\ & & & & & \ddots \\ & & & & & & 0 \end{vmatrix} \begin{vmatrix} 1 & l_{12} & \dots & l_{1n} \\ 0 & 1 & \dots & l_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{vmatrix}, \quad (44)$$

where

$$f_{gk} = \frac{A \begin{pmatrix} 1 & 2 & \dots & k-1 & g \\ 1 & 2 & \dots & k-1 & k \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}}, \quad l_{kg} = \frac{A \begin{pmatrix} 1 & 2 & \dots & k-1 & k \\ 1 & 2 & \dots & k-1 & g \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}}, \quad (45)$$

$$(g = k+1, \dots, n; k = 1, 2, \dots, r),$$

and  $f_{gk}$  and  $l_{kg}$  are arbitrary for  $g = k+1, \dots, n; k = r+1, \dots, n$ .

3. The elimination method of Gauss, when applied to a matrix  $A = \|a_{ik}\|_1^n$  of rank  $r$  for which  $D_k \neq 0$  ( $k = 1, 2, \dots, r$ ), yields two matrices: a lower triangular matrix  $W$  with diagonal elements 1 and an upper triangular matrix  $G$  in which the first  $r$  diagonal elements are  $D_1, \frac{D_2}{D_1}, \dots, \frac{D_r}{D_{r-1}}$  and the last  $n - r$  rows consist entirely of zeros.  $G$  is the *Gaussian form* of the matrix  $A$ ;  $W$  is the transforming matrix.

For actual computation of the elements of  $W$  we recommend the following device.

We obtain the matrix  $W$  when we apply to the unit matrix  $E$  all the transformations (given by  $W_1, \dots, W_N$ ) that we have performed on  $A$  in the algorithm of Gauss (in this case we shall have instead of the product  $WA$ , equal to  $G$ , the product  $WE$ , equal to  $W$ ). Let us, therefore, write the unit matrix  $E$  on the right of  $A$ :

$$\left\| \begin{array}{cccccc} a_{11} & \dots & a_{1n} & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \dots & a_{nn} & 0 & & 1 \end{array} \right\|. \quad (46)$$

By applying all the transformations of the algorithm of Gauss to this rectangular matrix we obtain a rectangular matrix consisting of the two square matrices  $G$  and  $W$ :

$$(G, W).$$

Thus, the application of Gauss's algorithm to the matrix (46) gives the matrices  $G$  and  $W$  simultaneously.

If  $A$  is non-singular, so that  $|A| \neq 0$ , then  $|G| \neq 0$  as well. In this case, (33) implies that  $A^{-1} = G^{-1}W$ . Since  $G$  and  $W$  are determined by means of the algorithm of Gauss, the task of finding the inverse matrix  $A^{-1}$  reduces to determining  $G^{-1}$  and multiplying  $G^{-1}$  by  $W$ .

Although there is no difficulty in finding the inverse matrix  $G^{-1}$  once the matrix  $G$  has been determined, because  $G$  is triangular, the operations involved can nevertheless be avoided. For this purpose we introduce, together with the matrices  $G$  and  $W$ , similar matrices  $G_1$  and  $W_1$  for the transposed matrix  $A^T$ . Then  $A^T = W_1^{-1}G_1$ , i.e.,

$$A = G_1^T W_1^{T-1}. \quad (47)$$

Let us compare (33') with (44):

$$A = W^{-1}G, \quad A = FDL.$$

These equations may be regarded as two distinct decompositions of the form (35); here we take the product  $DL$  as the second factor  $C$ . Since the first  $r$  diagonal elements of the first factors are the same (they are equal to 1), their first  $r$  columns coincide. But then, since the last  $n - r$  columns of  $F$  may be chosen arbitrarily, we chose them such that

$$F = W^{-1}. \quad (48)$$

On the other hand, a comparison of (47) with (44),

$$A = G_1^T W_1^T, \quad A = FDL,$$

shows that we may also select the arbitrary elements of  $L$  in such a way that

$$L = W_1^{-1}. \quad (49)$$

Replacing  $F$  and  $L$  in (44) by their expressions (48) and (49), we obtain

$$A = W^{-1} D W_1^T. \quad (50)$$

Comparing this equation with (33') and (47) we find:

$$G = D W_1^T, \quad G_1^T = W^{-1} D. \quad (51)$$

We now introduce the diagonal matrix

$$\hat{D} = \left\{ \frac{1}{D_1}, \frac{D_1}{D_2}, \dots, \frac{D_{r-1}}{D_r}, 0, \dots, 0 \right\}. \quad (52)$$

Since

$$D = D \hat{D} D,$$

it follows from (50) and (51) that

$$A = G_1^T \hat{D} G. \quad (53)$$

Formula (53) shows that the decomposition of  $A$  into triangular factors can be obtained by applying the algorithm of Gauss to the matrices  $A$  and  $A^T$ .

Now let  $A$  be non-singular ( $r = n$ ). Then  $|D| \neq 0$ ,  $\hat{D} = D^{-1}$ . Therefore it follows from (50) that

$$A^{-1} = W_1^T \hat{D} W. \quad (54)$$

This formula yields an effective computation of the inverse matrix  $A^{-1}$  by the application of Gauss's algorithm to the rectangular matrices

$$(A, E) \quad (A^T, E).$$

If, in particular, we take as our  $A$  a symmetrical matrix  $S$ , then  $G_1$  coincides with  $G$  and  $W_1$  with  $W$ , and therefore formulas (53) and (54) assume the form

$$S = G^T \hat{D} G, \quad (55)$$

$$S^{-1} = W^T \hat{D} W. \quad (56)$$

### § 5. The Partition of a Matrix into Blocks. The Technique of Operating with Partitioned Matrices. The Generalized Algorithm of Gauss

It often becomes necessary to use matrices that are partitioned into rectangular parts—'cells' or 'blocks.' In the present section we deal with such partitioned matrices.

1. Let a rectangular matrix

$$A = \|a_{ik}\| \quad (i = 1, 2, \dots, m; k = 1, 2, \dots, n) \quad (57)$$

be given.

By means of horizontal and vertical lines we dissect  $A$  into rectangular blocks:

$$A = \begin{pmatrix} \overset{n_1}{\widehat{A}_{11}} & \overset{n_2}{\widehat{A}_{12}} & \dots & \overset{n_t}{\widehat{A}_{1t}} \\ \widehat{A}_{21} & \widehat{A}_{22} & \dots & \widehat{A}_{2t} \\ \dots & \dots & \dots & \dots \\ \widehat{A}_{s1} & \widehat{A}_{s2} & \dots & \widehat{A}_{st} \end{pmatrix} \begin{matrix} \} m_1 \\ \} m_2 \\ \vdots \\ \} m_s \end{matrix}. \quad (58)$$

We shall say of matrix (58) that it is *partitioned* into  $st$  blocks, or *cells*  $\widehat{A}_{\alpha\beta}$  of dimensions  $m_\alpha \times n_\beta$  ( $\alpha = 1, 2, \dots, s$ ;  $\beta = 1, 2, \dots, t$ ), or that it is represented in the form of a *partitioned*, or *blocked*, matrix. Instead of (58) we shall simply write

$$A = (A_{\alpha\beta}) \quad (\alpha = 1, 2, \dots, s; \beta = 1, 2, \dots, t). \quad (59)$$

In the case  $s = t$  we shall use the following notation:

$$A = (A_{\alpha\beta})_1^1. \quad (60)$$

Operations on partitioned matrices are performed according to the same formal rules as in the case in which we have numerical elements instead of blocks. For example, let  $A$  and  $B$  be two rectangular matrices of equal dimensions partitioned into blocks in exactly the same way:

$$A = (A_{\alpha\beta}), \quad B = (B_{\alpha\beta}) \quad (\alpha = 1, 2, \dots, s; \beta = 1, 2, \dots, t) \quad (61)$$

It is easy to verify that

$$A + B = (A_{\alpha\beta} + B_{\alpha\beta}) \quad (\alpha = 1, 2, \dots, s; \beta = 1, 2, \dots, t). \quad (62)$$

We have to consider multiplication of partitioned matrices in more detail. We know (see Chapter I, p. 6) that for the multiplication of two rectangular matrices  $A$  and  $B$  the length of the rows of the first factor  $A$  must be the same as the height of the columns of the second factor  $B$ . For 'block' multiplication of these matrices we require, in addition, that the partitioning into blocks be such that the horizontal dimensions in the first factor are the same as the corresponding vertical dimensions in the second:

$$A = \begin{pmatrix} \begin{matrix} \overline{A}_{11} & \overline{A}_{12} & \dots & \overline{A}_{1t} \end{matrix} \\ \begin{matrix} \overline{A}_{21} & \overline{A}_{22} & \dots & \overline{A}_{2t} \end{matrix} \\ \vdots \\ \begin{matrix} \overline{A}_{s1} & \overline{A}_{s2} & \dots & \overline{A}_{st} \end{matrix} \end{pmatrix} \begin{matrix} \left. \begin{matrix} m_1 \\ m_2 \\ \vdots \\ m_s \end{matrix} \right\} \\ \left. \begin{matrix} m_1 \\ m_2 \\ \vdots \\ m_s \end{matrix} \right\} \end{matrix}, \quad B = \begin{pmatrix} \begin{matrix} \overline{B}_{11} & \overline{B}_{12} & \dots & \overline{B}_{1u} \end{matrix} \\ \begin{matrix} \overline{B}_{21} & \overline{B}_{22} & \dots & \overline{B}_{2u} \end{matrix} \\ \vdots \\ \begin{matrix} \overline{B}_{t1} & \overline{B}_{t2} & \dots & \overline{B}_{tu} \end{matrix} \end{pmatrix} \begin{matrix} \left. \begin{matrix} n_1 \\ n_2 \\ \vdots \\ n_t \end{matrix} \right\} \\ \left. \begin{matrix} n_1 \\ n_2 \\ \vdots \\ n_t \end{matrix} \right\} \end{matrix}. \quad (63)$$

Then it is easy to verify that

$$AB = C = (C_{\alpha\beta}), \quad \text{where} \quad C_{\alpha\beta} = \sum_{\delta=1}^t A_{\alpha\delta} B_{\delta\beta} \quad \begin{pmatrix} \alpha = 1, 2, \dots, s \\ \beta = 1, 2, \dots, u \end{pmatrix}. \quad (64)$$

We mention separately the special case in which one of the factors is a *quasi-diagonal* matrix. Let  $A$  be quasi-diagonal, i.e., let  $s = t$  and  $A_{\alpha\beta} = 0$  for  $\alpha \neq \beta$ . In this case formula (64) gives

$$C_{\alpha\beta} = A_{\alpha\alpha} B_{\alpha\beta} \quad (\alpha = 1, 2, \dots, s; \beta = 1, 2, \dots, u). \quad (65)$$

When a partitioned matrix is multiplied on the left by a quasi-diagonal matrix, then the rows of the matrix are multiplied on the left by the corresponding diagonal blocks of the quasi-diagonal matrix.

Now let  $B$  be a quasi-diagonal matrix, i.e., let  $t = u$  and  $B_{\alpha\beta} = 0$  for  $\alpha \neq \beta$ . Then we obtain from (64):

$$C_{\alpha\beta} = A_{\alpha\beta} B_{\beta\beta} \quad (\alpha = 1, 2, \dots, s; \beta = 1, 2, \dots, u). \quad (66)$$

When a partitioned matrix is multiplied on the right by a quasi-diagonal matrix, then all the columns of the partitioned matrix are multiplied on the right by the corresponding diagonal cells of the quasi-diagonal matrix.

Note that the multiplication of square partitioned matrices of one and the same order is always feasible if the factors are split into equal quadratic schemes of blocks and there are square matrices on the diagonal places in each factor.

The partitioned matrix (58) is called *upper (lower) quasi-triangular* if  $s = t$  and all  $A_{\alpha\beta} = O$  for  $\alpha > \beta$  ( $\alpha < \beta$ ). A quasi-diagonal matrix is a special case of a quasi-triangular matrix.

From the formulas (64) it is easy to see that:

The product of two upper (lower) quasi-triangular matrices is itself an upper (lower) quasi-triangular matrix;<sup>6</sup> the diagonal cells of the product are obtained by multiplying the corresponding diagonal cells of the factors.

For when we set  $s = t$  in (64) and

$$A_{\alpha\beta} = O, \quad B_{\alpha\beta} = O \quad \text{for } \alpha < \beta,$$

we find

$$\text{and} \quad \left. \begin{aligned} C_{\alpha\beta} &= O & \text{for } \alpha < \beta \\ C_{\alpha\alpha} &= A_{\alpha\alpha} B_{\alpha\alpha} \end{aligned} \right\} \quad (\alpha, \beta = 1, 2, \dots, s).$$

The case of lower quasi-triangular matrices is treated similarly.

We mention a rule for the calculation of the determinant of a quasi-triangular matrix. This rule can be obtained from the Laplace expansion.

If  $A$  is a quasi-triangular matrix (in particular, a quasi-diagonal matrix), then the determinant of the matrix is equal to the product of the determinant of the diagonal cells:

$$|A| = |A_{11}| |A_{22}| \cdots |A_{ss}|. \quad (67)$$

## 2. Let a partitioned matrix

$$A = \begin{pmatrix} \overset{n_1}{\widetilde{A}_{11}} & \overset{n_2}{\widetilde{A}_{12}} & \cdots & \overset{n_s}{\widetilde{A}_{1s}} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ A_{s1} & A_{s2} & \cdots & A_{ss} \end{pmatrix} \begin{Bmatrix} m_1 \\ m_2 \\ \vdots \\ m_s \end{Bmatrix}. \quad (68)$$

<sup>6</sup> It is assumed here that the block multiplication is feasible.

be given. To the  $\alpha$ -th row of submatrices we add the  $\beta$ -th row, multiplied on the left by a rectangular matrix  $X$  of dimension  $m_\alpha \times n_\beta$ . We obtain a partitioned matrix

$$B = \begin{pmatrix} A_{11} & \dots & A_{1t} \\ \dots & \dots & \dots \\ A_{\alpha 1} + XA_{\beta 1} & \dots & A_{\alpha t} + XA_{\beta t} \\ \dots & \dots & \dots \\ A_{\beta 1} & \dots & A_{\beta t} \\ \dots & \dots & \dots \\ A_{s1} & \dots & A_{st} \end{pmatrix}. \quad (69)$$

We introduce an auxiliary square matrix  $V$ , which we give in the form of a square scheme of blocks:

$$V = \begin{pmatrix} \overbrace{E}^{m_1} & \dots & \overbrace{O}^{m_\alpha} & \dots & \overbrace{O}^{m_\beta} & \dots & \overbrace{O}^{m_s} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ O & \dots & E & \dots & X & \dots & O \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ O & \dots & O & \dots & E & \dots & O \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ O & \dots & O & \dots & O & \dots & E \end{pmatrix} \begin{matrix} \} m_1 \\ \\ \} m_\alpha \\ \\ \} m_\beta \\ \\ \} m_s \end{matrix} \quad (70)$$

In the diagonal blocks of  $V$  there are unit matrices of order  $m_1, m_2, \dots, m_s$ , respectively; all the non-diagonal blocks of  $V$  are equal to zero except the block  $X$  that lies at the intersection of the  $\alpha$ -th row and  $\beta$ -th column.

It is easy to see that

$$VA = B. \quad (71)$$

As  $V$  is non-singular, we have<sup>7</sup> for the ranks of  $A$  and  $B$ :

$$r_A = r_B. \quad (72)$$

In the special case where  $A$  is a square matrix, we have from (70):

$$|V||A| = |B|. \quad (73)$$

But the determinant of the quasi-triangular matrix  $V$  is 1:

$$|V| = 1. \quad (74)$$

Hence

$$|A| = |B|. \quad (75)$$

<sup>7</sup> See p. 12.

The same conclusion holds when we add to an arbitrary column of (68) another column multiplied on the right by a rectangular matrix  $x$  of suitable dimensions.

The results obtained can be formulated as the following theorem.

**THEOREM 3:** *If to the  $\alpha$ -th row (column) of the blocks of the partitioned matrix  $A$  we add the  $\beta$ -th row (column) multiplied on the left (right) by a rectangular matrix  $X$  of the corresponding dimensions, then the rank of  $A$  remains unchanged under this transformation and, if  $A$  is a square matrix, the determinant of  $A$  is also unchanged.*

**3.** We now consider the special case in which the diagonal block  $A_{11}$  in  $A$  is square and non-singular ( $|A_{11}| \neq 0$ ).

To the  $\alpha$ -th row of  $A$  we add the first row multiplied on the left by ( $\alpha = 2, \dots, s$ ). We thus obtain the matrix

$$B_1 = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1t} \\ 0 & A_{22}^{(1)} & \dots & A_{2t}^{(1)} \\ \dots & \dots & \dots & \dots \\ 0 & A_{s2}^{(1)} & \dots & A_{st}^{(1)} \end{pmatrix}, \quad (76)$$

where

$$A_{\alpha\beta}^{(1)} = -A_{\alpha 1} A_{11}^{-1} A_{1\beta} + A_{\alpha\beta} \quad (\alpha = 2, \dots, s; \beta = 2, \dots, t). \quad (77)$$

If the matrix  $A_{22}^{(1)}$  is square and non-singular, then the process can be continued. In this way we arrive at the *generalized algorithm of Gauss*.

Let  $A$  be a square matrix. Then

$$|A| = |B_1| = |A_{11}| \begin{vmatrix} A_{22}^{(1)} & \dots & A_{2t}^{(1)} \\ \dots & \dots & \dots \\ A_{s2}^{(1)} & \dots & A_{st}^{(1)} \end{vmatrix}. \quad (78)$$

Formula (78) reduces the computation of the determinant  $|A|$ , consisting of  $st$  blocks to the computation of a determinant of lower order consisting of  $(s-1) \cdot (t-1)$  blocks.<sup>s</sup>

Let us consider a determinant  $A$  partitioned into four blocks:

$$A = \begin{vmatrix} A & B \\ C & D \end{vmatrix}. \quad (79)$$

where  $A$  and  $D$  are square matrices.

Suppose  $|A| \neq 0$ . Then from the second row we subtract the first multiplied on the left by  $CA^{-1}$ . We obtain

<sup>s</sup> If  $A_{22}^{(1)}$  is a square matrix and  $|A_{22}^{(1)}| \neq 0$ , then this determinant of  $(s-1) \cdot (t-1)$  blocks can again be subjected to such a transformation, etc.



$$\Delta = \begin{vmatrix} A & B \\ O & D - CA^{-1}B \end{vmatrix} = |A| |D - CA^{-1}B|. \quad (\text{I})$$

Similarly, if  $|D| \neq 0$ , we subtract from the first row in  $\Delta$  the second multiplied on the left by  $BD^{-1}$ , obtaining

$$\Delta = \begin{vmatrix} A - BD^{-1}C & O \\ C & D \end{vmatrix} = |A - BD^{-1}C| |D|. \quad (\text{II})$$

In the special case in which all four matrices  $A, B, C, D$  are square (of one and the same order  $n$ ), we deduce from (I) and (II) the *formulas of Schur*, which reduce the computation of a determinant of order  $2n$  to the computation of a determinant of order  $n$ :

$$\Delta = |AD - ACA^{-1}B| \quad (A \neq 0), \quad (\text{Ia})$$

$$\Delta = |AD - BD^{-1}CD| \quad (D \neq 0). \quad (\text{IIa})$$

If the matrices  $A$  and  $C$  are permutable, then it follows from (Ia) that

$$\Delta = |AD - CB| \quad (\text{provided } AC = CA). \quad (\text{Ib})$$

Similarly, if  $C$  and  $D$  are permutable, then

$$\Delta = |AD - BC| \quad (\text{provided } CD = DC). \quad (\text{IIb})$$

Formula (Ib) was obtained under the assumption  $|A| \neq 0$ , and (IIb) under the assumption  $|D| \neq 0$ . However, these restrictions can be removed by continuity arguments.

From formulas (I)-(IIb) we can obtain another six formulas by replacing  $A$  and  $D$  on the right-hand sides simultaneously by  $B$  and  $C$ .

*Example.*

$$\Delta = \begin{vmatrix} 1 & 0 & b_1 & b_2 \\ 0 & 1 & b_3 & b_4 \\ c_1 & c_2 & d_1 & d_2 \\ c_3 & c_4 & d_3 & d_4 \end{vmatrix}.$$

By formula (Ib),

$$\Delta = \begin{vmatrix} d_1 - c_1b_1 - c_2b_3 & d_2 - c_1b_2 - c_2b_4 \\ d_3 - c_3b_1 - c_4b_3 & d_4 - c_3b_2 - c_4b_4 \end{vmatrix}.$$

4. From Theorem 3 there follows also

**THEOREM 4:** *If a rectangular matrix  $R$  is represented in partitioned form*

$$R = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad (80)$$

*where  $A$  is a square non-singular matrix of order  $n$  ( $|A| \neq 0$ ), then the rank of  $R$  is equal to  $n$  if and only if*

$$D = CA^{-1}B. \quad (81)$$

*Proof.* We subtract from the second row of blocks of  $R$  the first, multiplied on the left by  $CA^{-1}$ . Then we obtain the matrix

$$T = \begin{pmatrix} A & B \\ O & D - CA^{-1}B \end{pmatrix}. \quad (82)$$

By Theorem 3, the matrices  $R$  and  $T$  have the same rank. But the rank of  $T$  coincides with the rank of  $A$  (namely,  $n$ ) if and only if  $D - CA^{-1}B = O$ , i.e., when (80) holds. This proves the theorem.

From Theorem 4 there follows an algorithm<sup>9</sup> for the construction of the inverse matrix  $A^{-1}$  and, more generally, the product  $CA^{-1}B$ , where  $B$  and  $C$  are rectangular matrices of dimensions  $n \times p$  and  $q \times n$ .

By means of Gauss's algorithm,<sup>10</sup> we reduce the matrix

$$\begin{pmatrix} A & B \\ -C & O \end{pmatrix} \quad (|A| \neq 0) \quad (83)$$

to the form

$$\begin{pmatrix} G & B_1 \\ O & X \end{pmatrix}. \quad (84)$$

We will show that

$$X = CA^{-1}B. \quad (85)$$

For, the same transformation that was applied to the matrix (83) reduces the matrix

<sup>9</sup> See [181].

<sup>10</sup> We do not apply here the entire algorithm of Gauss to the matrix (83) but only the first  $n$  steps of the algorithm, where  $n$  is the order of the matrix. This can be done if the conditions (15) hold for  $p = n$ . But if these conditions do not hold, then, since  $|A| \neq 0$ , we may renumber the first  $n$  rows (or the first  $n$  columns) of the matrix (83) so that the  $n$  steps of Gauss's algorithm turn out to be feasible. Such a modified Gaussian algorithm is sometimes applied even when the conditions (15), with  $p = n$ , are satisfied.

to the form

$$\begin{pmatrix} A & B \\ -C & -CA^{-1}B \end{pmatrix} \quad (86)$$

$$\begin{pmatrix} G & B_1 \\ O & X - CA^{-1}B \end{pmatrix}. \quad (87)$$

By Theorem 4, the matrix (86) is of rank  $n$  ( $n$  is the order of  $A$ ). But then (87) must also be of rank  $n$ . Hence  $X - CA^{-1}B = O$ , i.e., (85) holds

In particular, if  $B = y$ , where  $y$  is a column matrix, and  $C = E$ , then

$$X = A^{-1}y.$$

Therefore, when we apply Gauss's algorithm to the matrix

$$\begin{pmatrix} A & y \\ -E & O \end{pmatrix},$$

we obtain the solution of the system of equations

$$Ax = y.$$

Further, if in (83) we set  $B = C = E$ , then by applying the algorithm of Gauss to the matrix

$$\begin{pmatrix} A & E \\ -E & O \end{pmatrix},$$

we obtain

$$\begin{pmatrix} G & W \\ O & X \end{pmatrix},$$

where

$$X = A^{-1}.$$

Let us illustrate this method by finding  $A^{-1}$  in the following example.

*Example.* Let

$$A = \begin{vmatrix} 2 & 1 & 1 \\ 1 & 0 & 2 \\ 3 & 1 & 2 \end{vmatrix}.$$

It is required to compute  $A^{-1}$ .

We apply a somewhat modified elimination method<sup>11</sup> to the matrix

<sup>11</sup> See the preceding footnote.

$$\left\| \begin{array}{cccccc} 2 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 0 & 1 & 0 \\ 3 & 1 & 2 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{array} \right\|.$$

To all the rows we add certain multiples of the second row and we arrange that all the elements of the first column, except the second, become zero. Then we add to all the rows, except the second, the third row multiplied by certain factors and see to it that in the second column all the elements, except the second and third, become zero. Then we add to the last three rows the first row with suitable factors and obtain a matrix of the form

$$\left\| \begin{array}{cccccc} * & * & * & * & * & * \\ * & * & * & * & * & * \\ * & * & * & * & * & * \\ 0 & 0 & 0 & -2 & -1 & 2 \\ 0 & 0 & 0 & 4 & 1 & -3 \\ 0 & 0 & 0 & 1 & 1 & -1 \end{array} \right\|.$$

Therefore

$$A^{-1} = \left\| \begin{array}{ccc} -2 & -1 & 2 \\ 4 & 1 & -3 \\ 1 & 1 & -1 \end{array} \right\|.$$

## CHAPTER III

### LINEAR OPERATORS IN AN $n$ -DIMENSIONAL VECTOR SPACE

Matrices constitute the fundamental analytic apparatus for the study of linear operators in an  $n$ -dimensional space. The study of these operators, in turn, enables us to divide all matrices into classes and to exhibit the significant properties that all matrices of one and the same class have in common.

In the present chapter we shall expound the simpler properties of linear operators in an  $n$ -dimensional space. The investigation will be continued in Chapters VII and IX.

#### § 1. Vector Spaces

1. Let  $\mathbf{R}$  be a set of arbitrary elements  $\mathbf{x}, \mathbf{y}, \mathbf{z}, \dots$  in which two operations are defined:<sup>1</sup> the operation of 'addition' and the operation of 'multiplication by a number of the field  $\mathbf{r}$ .' We postulate that these operations can always be performed uniquely in  $\mathbf{R}$  and that the following rules hold for arbitrary elements  $\mathbf{x}, \mathbf{y}, \mathbf{z}$  of  $\mathbf{R}$  and numbers  $\alpha, \beta$  of  $\mathbf{r}$ :

1.  $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ .
2.  $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ .
3. There exists an element  $\mathbf{o}$  in  $\mathbf{R}$  such that the product of the number 0 with any element  $\mathbf{x}$  of  $\mathbf{R}$  is equal to  $\mathbf{o}$ :  
 $0 \cdot \mathbf{x} = \mathbf{o}$ .
4.  $1 \cdot \mathbf{x} = \mathbf{x}$ .
5.  $\alpha(\beta \mathbf{x}) = (\alpha\beta) \mathbf{x}$ .
6.  $(\alpha + \beta) \mathbf{x} = \alpha \mathbf{x} + \beta \mathbf{x}$ .
7.  $\alpha(\mathbf{x} + \mathbf{y}) = \alpha \mathbf{x} + \alpha \mathbf{y}$ .

---

<sup>1</sup> These operations will be denoted by the usual signs '+' and '·'; the latter sign will sometimes be omitted.

**DEFINITION 1:** A set  $\mathbf{R}$  of elements in which two operations—'addition of elements and 'multiplication of elements of  $\mathbf{R}$  by a number of  $\mathbb{F}$ '—can always be performed uniquely and for which postulates 1.-7. hold is called a *vector space* (over the field  $\mathbb{F}$ ) and the elements are called *vectors*.<sup>2</sup>

**DEFINITION 2.** The vectors  $\mathbf{x}, \mathbf{y}, \dots, \mathbf{u}$  of  $\mathbf{R}$ , are called *linearly dependent* if there exist numbers  $\alpha, \beta, \dots, \delta$  in  $\mathbb{F}$ , not all zero, such that

$$\alpha \mathbf{x} + \beta \mathbf{y} + \dots + \delta \mathbf{u} = \mathbf{o}. \quad (1)$$

If such a linear dependence does not hold, then the vectors  $\mathbf{x}, \mathbf{y}, \dots, \mathbf{u}$  are called *linearly independent*.

If the vectors  $\mathbf{x}, \mathbf{y}, \dots, \mathbf{u}$  are linearly dependent, then one of the vectors can be represented as a linear combination, with coefficients in  $\mathbb{F}$ , of the remaining ones. For example, if  $\alpha \neq 0$  in (1), then

$$\mathbf{x} = -\frac{\beta}{\alpha} \mathbf{y} - \dots - \frac{\delta}{\alpha} \mathbf{u}.$$

**DEFINITION 3.** The space  $\mathbf{R}$  is called *finite-dimensional* and the number  $n$  is called the *dimension* of the space if there exist  $n$  linearly independent vectors in  $\mathbf{R}$ , while any  $n+1$  vectors in  $\mathbf{R}$  are linearly dependent. If the space contains linearly independent systems of an arbitrary number of vectors, then it is called *infinite-dimensional*.

In this book we shall study mainly finite-dimensional spaces.

**DEFINITION 4.** A system of  $n$  linearly independent vectors  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  of an  $n$ -dimensional space, given in a definite order, is called a *basis* of the space.

**2. Example 1.** The set of all ordinary vectors (directed geometrical segments) is a three-dimensional vector space. The part of this space that consists of the vectors parallel to some plane is a two-dimensional space, and all the vectors parallel to a given line form a one-dimensional vector space.

**Example 2.** Let us call a column  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  of  $n$  numbers of  $\mathbb{F}$  a vector (where  $n$  is a fixed number). We define the basic operations as operations on column matrices:

<sup>2</sup> It is easy to see that all the usual properties of the operations of addition and of multiplication by a number follow from properties 1.-7. For example, for arbitrary  $\mathbf{x}$  of  $\mathbf{R}$  we have:

$$\begin{aligned} \mathbf{x} + \mathbf{o} &= \mathbf{x} \quad [\mathbf{x} + \mathbf{o} = 1 \cdot \mathbf{x} + 0 \cdot \mathbf{x} = (1+0) \cdot \mathbf{x} = 1 \cdot \mathbf{x} = \mathbf{x}]; \\ \mathbf{x} + (-\mathbf{x}) &= \mathbf{o}, \text{ where } -\mathbf{x} = (-1) \cdot \mathbf{x}; \end{aligned}$$

etc.

$$(x_1, x_2, \dots, x_n) + (y_1, y_2, \dots, y_n) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n),$$

$$\alpha (x_1, x_2, \dots, x_n) = (\alpha x_1, \alpha x_2, \dots, \alpha x_n).$$

The null vector is the column  $(0, 0, \dots, 0)$ . It is easy to verify that all the postulates 1-7. are satisfied. The vectors form an  $n$ -dimensional space. As a basis of the space we can take, for example, the column of unit matrices of order  $n$ :

$$(1, 0, \dots, 0), (0, 1, \dots, 0), \dots, (0, 0, \dots, 1).$$

The space thus defined is often called the  $n$ -dimensional number space.

*Example 3.* The set of all infinite sequences  $(x_1, x_2, \dots, x_n, \dots)$  in which the operations are defined in a natural way, i.e.,

$$(x_1, x_2, \dots, x_n, \dots) + (y_1, y_2, \dots, y_n, \dots) = (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n, \dots),$$

$$\alpha (x_1, x_2, \dots, x_n, \dots) = (\alpha x_1, \alpha x_2, \dots, \alpha x_n, \dots),$$

is an infinite-dimensional space.

*Example 4.* The set of polynomials  $a_0 + a_1 t + \dots + a_{n-1} t^{n-1}$  of degree  $< n$  with coefficients in  $\mathbb{F}$  is an  $n$ -dimensional vector space.<sup>3</sup> As a basis of this space we can take, say, the system of powers  $t^0, t^1, \dots, t^{n-1}$ .

The set of all such polynomials (without a bound on the degree) form an infinite-dimensional space.

*Example 5.* The set of all functions defined on a closed interval  $[a, b]$  form an infinite-dimensional space.

3. Let the vectors  $e_1, e_2, \dots, e_n$  form a basis of an  $n$ -dimensional vector space  $\mathbf{R}$  and let  $\mathbf{x}$  be an arbitrary vector of the space. Then the vectors  $\mathbf{x}, e_1, e_2, \dots, e_n$  are linearly dependent (because there are  $n + 1$  of them):

$$\alpha_0 \mathbf{x} + \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n = \mathbf{0},$$

where at least one of the numbers  $\alpha_0, \alpha_1, \dots, \alpha_n$  is different from zero. But in this case we must have  $\alpha_0 \neq 0$ , since the vectors  $e_1, e_2, \dots, e_n$  cannot be linearly dependent. Therefore

$$\mathbf{x} = x_1 e_1 + x_2 e_2 + \dots + x_n e_n \quad (2)$$

where  $x_i = -\alpha_i/\alpha_0$  ( $i = 1, 2, \dots, n$ ).

Note that the numbers  $x_1, x_2, \dots, x_n$  are uniquely determined when the vector  $\mathbf{x}$  and the basis  $e_1, e_2, \dots, e_n$  are given. For if there is another decomposition of  $\mathbf{x}$  besides (2),

$$\mathbf{x} = x'_1 e_1 + x'_2 e_2 + \dots + x'_n e_n, \quad (3)$$

<sup>3</sup> The basic operations are taken to be ordinary addition of polynomials and multiplication of a polynomial by a number.





Thus, the following theorem holds:

**THEOREM 1:** *In order that the vectors  $x_1, x_2, \dots, x_m$  be linearly independent it is necessary and sufficient that the rank  $r$  of the matrix formed from the coordinates of these vectors in an arbitrary basis*

$$\begin{vmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{vmatrix} \quad (7)$$

be equal to  $m$ , i.e., to the number of vectors.

*Note.* The linear independence of the vectors  $x_1, x_2, \dots, x_m$  means that the columns of the matrix (7) are linearly independent, since the  $k$ -th column consists of the coordinates of  $x_k$  ( $k = 1, 2, \dots, m$ ). By the theorem, therefore, if the columns of a matrix are linearly independent, then the rank of the matrix is equal to the number of columns. Hence it follows that in an arbitrary rectangular matrix the maximal number of linearly independent columns is equal to the rank of the matrix. Moreover, if we transpose the matrix, i.e., change the rows into columns and the columns into rows, then the rank obviously remains unchanged. Hence in a rectangular matrix the number of linearly independent columns is always equal to the number of linearly independent rows and equal to the rank of the matrix.<sup>4</sup>

**5.** If in an  $n$ -dimensional space a basis  $e_1, e_2, \dots, e_n$  has been chosen, then to every vector  $x$  there corresponds uniquely the column  $x = (x_1, x_2, \dots, x_n)$ , where  $x_1, x_2, \dots, x_n$  are the coordinates of  $x$  in the given basis. Thus, the choosing of a basis establishes a one-to-one correspondence between the vectors of an arbitrary  $n$ -dimensional vector space  $R$  and the vectors of the  $n$ -dimensional number space  $R'$  considered in Example 2. Here the sum of vectors in  $R$  corresponds to the sum of the corresponding vectors of  $R'$ . The analogous correspondence holds for the product of a vector by a number  $a$  of  $F$ . In other words, an arbitrary  $n$ -dimensional vector space is *isomorphic* to the  $n$ -dimensional number space, and therefore *all vector spaces of the same number  $n$  of dimensions over the same number field  $F$  are isomorphic*. This means that to within isomorphism there exists only one  $n$ -dimensional vector space for a given number field.

<sup>4</sup> This proposition follows from Theorem 1, in the proof of which we have started from the well-known property of a system of linear homogeneous equations: a non-zero solution exists only when the rank of the coefficient matrix is less than the number of unknowns. For a proof of Theorem 1 independent of this property, see § 5.

The reader may ask why we have introduced an 'abstract'  $n$ -dimensional space if it coincides to within isomorphism with the  $n$ -dimensional number space. Indeed, we could have defined a vector as a system of  $n$  numbers given in a definite order and could have introduced the operations on these vectors in the very way it was done in Example 2. But we would then have mixed up properties of vectors that do not depend on the choice of a basis with properties of a particular basis. For example, the fact that all the coordinates of a vector are zero is a property of the vector itself; it does not depend on the choice of basis. But the equality of all its coordinates is not a property of the vector itself, because it disappears under a change of basis. *The axiomatic definition of a vector space immediately singles out the properties of vectors that do not depend on the choice of a basis.*

## § 2. A Linear Operator Mapping an $n$ -Dimensional Space into an $m$ -Dimensional Space

1. We consider a linear transformation

$$\left. \begin{aligned} y_1 &= a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ y_2 &= a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ &\vdots \\ y_m &= a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{aligned} \right\} \quad (8)$$

whose coefficients belong to the number field  $\mathbb{F}$  as well as two vector spaces over  $\mathbb{F}$ : an  $n$ -dimensional space  $\mathbf{R}$  and an  $m$ -dimensional space  $\mathbf{S}$ . We choose a basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  in  $\mathbf{R}$  and a basis  $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_m$  in  $\mathbf{S}$ . Then the transformation (8) associates with every vector  $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$  of  $\mathbf{R}$  a certain vector  $\mathbf{y} = \sum_{k=1}^m y_k \mathbf{g}_k$  of  $\mathbf{S}$ , i.e., the transformation (8) determines a certain operator that sets up a correspondence between the vector  $\mathbf{x}$  and the vector  $\mathbf{y}$ .  $\mathbf{y} = \mathbf{A}\mathbf{x}$ . It is easy to see that this operator  $\mathbf{A}$  has the property of linearity, which we formulate as follows:

DEFINITION 5: An operator  $A$  mapping  $R$  into  $S$ , i.e., associating with every vector  $x$  of  $R$  a certain vector  $y = Ax$  of  $S$  is called linear if for arbitrary  $x, x_1$  of  $R$  and  $\alpha$  of  $F$

$$A(x_1 + x_2) = Ax_1 + Ax_2, \quad A(\alpha x_1) = \alpha Ax_1. \quad (9)$$

Thus, the transformation (8), for a given basis in  $\mathbf{R}$  and a given basis in  $\mathbf{S}$ , determines a linear operator mapping  $\mathbf{R}$  into  $\mathbf{S}$ .

We shall now show the converse, i.e., that for an arbitrary linear operator  $A$  mapping  $R$  into  $S$  and arbitrary bases  $e_1, e_2, \dots, e_n$  in  $R$  and  $g_1, g_2, \dots, g_m$  in  $S$ , there exists a rectangular matrix with elements in  $F$

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \quad (10)$$

such that the linear transformation (8) formed by means of this matrix expresses the coordinates of the transformed vector  $y = Ax$  in terms of the coordinates of the original vector  $x$ .

Let us, in fact, apply the operator  $A$  to the basis vector  $e_k$  and let the coordinates in the basis  $g_1, g_2, \dots, g_m$  of the vector  $Ae_k$  thus obtained be denoted by  $a_{1k}, a_{2k}, \dots, a_{mk}$  ( $k = 1, 2, \dots, n$ ):

$$Ae_k = \sum_{i=1}^m a_{ik} g_i \quad (k = 1, 2, \dots, n). \quad (11)$$

Multiplying both sides of (11) by  $x_k$  and summing from 1 to  $n$ , we obtain

$$\sum_{k=1}^n x_k Ae_k = \sum_{i=1}^m \left( \sum_{k=1}^n a_{ik} x_k \right) g_i;$$

hence

$$y = Ax = A \left( \sum_{k=1}^n x_k e_k \right) = \sum_{k=1}^n x_k Ae_k = \sum_{i=1}^m y_i g_i,$$

where

$$y_i = \sum_{k=1}^n a_{ik} x_k \quad (i = 1, 2, \dots, m),$$

and this is what we had to show.

Thus, for given bases of  $R$  and  $S$ : to every linear operator  $A$  mapping  $R$  into  $S$  there corresponds a rectangular matrix of dimension  $m \times n$  and, conversely, to every such matrix there corresponds a linear operator mapping  $R$  into  $S$ .

Here, in the matrix  $A$  corresponding to the operator  $A$ , the  $k$ -th column consists of the coordinates of the vector  $Ae_k$  ( $k = 1, 2, \dots, n$ ).

We denote by  $x = (x_1, x_2, \dots, x_n)$  and  $y = (y_1, y_2, \dots, y_m)$  the coordinate columns of the vectors  $x$  and  $y$ . Then the vector equation

$$y = Ax$$

corresponds to the matrix equation

$$y = Ax,$$

which is the matrix form of the transformation (8).

*Example.* We consider the set of all polynomials in  $t$  of degree  $\leq n-1$  with coefficients in  $F$ . This set forms an  $n$ -dimensional vector space  $R_n$  (see Example 4, p. 52). Similarly, the polynomials in  $t$  of degree  $\leq n-2$  with coefficients in  $F$  form a space  $R_{n-1}$ . The differentiation operator  $\frac{d}{dt}$  associates with every polynomial of  $R_n$  a certain polynomial in  $R_{n-1}$ . Thus, this operator maps  $R_n$  into  $R_{n-1}$ . The differentiation operator is linear, since

$$\frac{d}{dt} [\varphi(t) + \psi(t)] = \frac{d\varphi(t)}{dt} + \frac{d\psi(t)}{dt}, \quad \frac{d}{dt} [\alpha\varphi(t)] = \alpha \frac{d\varphi(t)}{dt}.$$

In  $R_n$  and  $R_{n-1}$ , we choose bases consisting of powers of  $t$ :

$$t^0 = 1, t, \dots, t^{n-1} \quad \text{and} \quad t^0 = 1, t, \dots, t^{n-2}.$$

Using formulas (11), we construct the rectangular matrix of dimension  $(n-1 \times n)$  corresponding to the differentiation operator  $\frac{d}{dt}$  in these bases:

$$\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & n-1 \end{pmatrix}$$

### § 3. Addition and Multiplication of Linear Operators

1. Let  $A$  and  $B$  be two linear operators mapping  $R$  into  $S$  and let the corresponding matrices be

$$A = \|a_{ik}\|, \quad B = \|b_{ik}\| \quad (i=1, 2, \dots, m; k=1, 2, \dots, n).$$

DEFINITION 6: The sum of the operators  $A$  and  $B$  is the operator  $C$  defined by the equation<sup>5</sup>

$$Cx = Ax + Bx \quad (x \in R). \quad (12)$$

On the basis of this definition it is easy to verify that the sum  $C = A + B$  of the linear operators  $A$  and  $B$  is itself a linear operator. Furthermore,

$$Ce_k = Ae_k + Be_k = \sum_{i=1}^n (a_{ik} + b_{ik}) e_k.$$

<sup>5</sup>  $x \in R$  means that the element  $x$  belongs to the set  $R$ . It is assumed that (12) holds for arbitrary  $x$  in  $R$ .

Hence it follows that the operator  $C$  corresponds to the matrix  $C = \|c_{ik}\|$ , where  $c_{ik} = a_{ik} + b_{ik}$  ( $i = 1, 2, \dots, m$ ;  $k = 1, 2, \dots, n$ ), i.e., the operator  $C$  corresponds to the matrix

$$C = A + B. \quad (13)$$

We would come to the same conclusion starting from the matrix equation

$$Cx = Ax + Bx \quad (14)$$

( $x$  is the coordinate column of the vector  $x$ ) corresponding to the vector equation (12). Since  $x$  is an arbitrary column, (13) follows from (14).

2. Let  $R$ ,  $S$ , and  $T$  be three vector spaces of dimension  $q$ ,  $n$ , and  $m$ , and let  $A$  and  $B$  be two linear operators, of which  $B$  maps  $R$  into  $S$  and  $A$  maps  $S$  into  $T$ ; in symbols:

$$R \xrightarrow{B} S \xrightarrow{A} T.$$

DEFINITION 7. The product of the operators  $A$  and  $B$  is the operator  $C$  for which

$$Cx = A(Bx) \quad (x \in R). \quad (15)$$

holds for every  $x$  of  $R$ .

The operator  $C$  maps  $R$  into  $T$ :

$$R \xrightarrow{C=AB} T.$$

From the linearity of the operators  $A$  and  $B$  follows the linearity of  $C$ . We choose arbitrary bases in  $R$ ,  $S$ , and  $T$  and denote by  $A$ ,  $B$ , and  $C$  the matrices corresponding, in this choice of basis, to the operators  $A$ ,  $B$ , and  $C$ . Then the vector equations

$$z = Ay, y = Bx, z = Cx \quad (16)$$

correspond to the matrix equations:

$$z = Ay, y = Bx, z = Cx,$$

where  $x$ ,  $y$ ,  $z$  are the coordinate columns of the vectors  $x$ ,  $y$ ,  $z$ . Hence

$$Cx = A(Bx) = (AB)x$$

and as the column  $x$  is arbitrary

$$C = AB. \quad (17)$$

Thus, the product  $C = AB$  of the operators  $A$  and  $B$  corresponds to the matrix  $C = \|c_{ij}\|$  ( $i = 1, 2, \dots, m$ ;  $j = 1, 2, \dots, q$ ), which is the product of the matrices  $A$  and  $B$ .





### § 5. Equivalent Matrices. The Rank of an Operator. Sylvester's Inequality

1. Let  $R$  and  $S$  be two vector spaces of dimension  $n$  and  $m$ , respectively, over the number field  $F$  and let  $A$  be a linear operator mapping  $R$  into  $S$ . In the present section we shall make clear how the matrix  $A$  corresponding to the given linear operator  $A$  changes when the bases in  $R$  and  $S$  are changed.

We choose arbitrary bases  $e_1, e_2, \dots, e_n$  in  $R$  and  $g_1, g_2, \dots, g_m$  in  $S$ . In these bases the operator  $A$  corresponds to a matrix  $A = \|a_{ik}\|$  ( $i = 1, 2, 3, \dots, m; k = 1, 2, \dots, n$ ). To the vector equation,

$$y = Ax \quad (26)$$

there corresponds the matrix equation

$$y = Ax, \quad (27)$$

where  $x$  and  $y$  are the coordinate columns for the vectors  $x$  and  $y$  in the bases  $e_1, e_2, \dots, e_n$  and  $g_1, g_2, \dots, g_m$ .

We now choose other bases  $e_1^*, e_2^*, \dots, e_n^*$  and  $g_1^*, g_2^*, \dots, g_m^*$  in  $R$  and  $S$ . In the new bases we shall have  $x^*, y^*, A^*$  instead of  $x, y, A$ . Here

$$y^* = A^* x^*. \quad (28)$$

Let us denote by  $Q$  and  $N$  the non-singular square matrices of order  $n$  and  $m$ , respectively, that realize the coordinate transformations in the spaces  $R$  and  $S$  on transition from the old bases to the new ones (see § 4):

$$x = Qx^*, \quad y = Ny^*. \quad (29)$$

Then we obtain from (27) and (29):

$$y^* = N^{-1}y = N^{-1}Ax = N^{-1}AQx^*. \quad (30)$$

Setting  $P = N^{-1}$ , we find from (28) and (30):

$$A^* = PAQ. \quad (31)$$

**DEFINITION 8:** Two rectangular matrices  $A$  and  $B$  of the same dimension are called equivalent if there exist two non-singular matrices  $P$  and  $Q$  such that<sup>8</sup>

$$B = PAQ. \quad (32)$$

<sup>8</sup> If the matrices  $A$  and  $B$  are of dimension  $m \times n$ , then in (32) the square matrix  $P$  is of order  $m$ , and  $Q$  of order  $n$ . If the elements of the equivalent matrices  $A$  and  $B$  belong to some number field, then  $P$  and  $Q$  may be chosen such that their elements belong to the



From (31) it follows that two matrices corresponding to one and the same linear operator  $A$  for different choices of bases in  $R$  and  $S$  are always equivalent. It is easy to see that, conversely, if a matrix  $A$  corresponds to the operator  $A$  for certain bases in  $R$  and  $S$ , and if a matrix  $B$  is equivalent to  $A$ , then it corresponds to the same linear operator for certain other bases in  $R$  and  $S$ .

Thus, to every linear operator mapping  $R$  into  $S$  there corresponds a class of equivalent matrices with elements in  $F$ .

2. The following theorem establishes a criterion for the equivalence of two matrices:

**THEOREM 2:** *Two rectangular matrices of the same dimension are equivalent if and only if they have the same rank.*

*Proof.* The condition is *necessary*. When a rectangular matrix is multiplied by an arbitrary non-singular square matrix (on the right or left), then its rank does not change (see Chapter I, p. 17). Therefore it follows from (32) that

$$r_A = r_B.$$

The condition is *sufficient*. Let  $A$  be a rectangular matrix of dimension  $\times n$ . It determines a linear operator  $A$  mapping the space  $R$  with the basis  $e_1, e_2, \dots, e_n$  into the space  $S$  with the basis  $g_1, g_2, \dots, g_m$ . Let  $r$  denote the number of linearly independent vectors among the vectors  $Ae_1, Ae_2, \dots, Ae_n$ . Without loss of generality we may assume that the vectors  $Ae_1, Ae_2, \dots, Ae_r$  are linearly independent<sup>9</sup> and that the remaining  $Ae_{r+1}, Ae_{r+2}, \dots, Ae_n$  are expressed linearly in terms of them:

$$Ae_k = \sum_{j=1}^r c_{kj} Ae_j \quad (k=r+1, \dots, n). \quad (33)$$

We define a new basis in  $R$  as follows:

$$e_i^* = \begin{cases} e_i & (i=1, 2, \dots, r), \\ e_i - \sum_{j=1}^r c_{ij} e_j & (i=r+1, \dots, n). \end{cases} \quad (34)$$

Then by (33),

$$Ae_k^* = 0 \quad (k=r+1, \dots, n). \quad (35)$$

Next, we set

$$Ae_j^* = g_j^* \quad (j=1, 2, \dots, r). \quad (36)$$

<sup>9</sup> This can be achieved by a suitable numbering of the basis vectors  $e_1, e_2, \dots, e_n$ .

The vectors  $\mathbf{g}_1^*, \mathbf{g}_2^*, \dots, \mathbf{g}_r^*$  are linearly independent. We supplement them with suitable vectors  $\mathbf{g}_{r+1}^*, \mathbf{g}_{r+2}^*, \dots, \mathbf{g}_m^*$  to obtain a basis  $\mathbf{g}_1^*, \mathbf{g}_2^*, \dots, \mathbf{g}_m^*$  of  $S$ .

The matrix corresponding to the same operator  $A$  in the new bases  $\mathbf{e}_1^*, \mathbf{e}_2^*, \dots, \mathbf{e}_n^*, \mathbf{g}_1^*, \mathbf{g}_2^*, \dots, \mathbf{g}_m^*$  has now, by (35) and (36), the form

$$I_r = \left\| \begin{array}{cccccc} \overbrace{1 \ 0 \ \dots \ 0}^r & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{array} \right\|. \quad (37)$$

Along the main diagonal of  $I_r$ , starting at the top, there are  $r$  units; all the remaining elements of  $I_r$  are zeros. Since the matrices  $A$  and  $I_r$  correspond to one and the same operator  $A$ , they are equivalent. As we have proved, equivalent matrices have the same rank. Hence the rank of the original matrix  $A$  is  $r$ .

We have shown that an arbitrary rectangular matrix of rank  $r$  is equivalent to the 'canonical' matrix  $I_r$ . But  $I_r$  is completely determined by specifying its dimensions  $m \times n$  and the number  $r$ . Therefore all rectangular matrices of given dimension  $m \times n$  and of given rank  $r$  are equivalent to one and the same matrix  $I_r$ , and consequently to each other. This completes the proof of the theorem.

**3.** Let  $A$  be a linear operator mapping an  $n$ -dimensional space  $R$  into an  $n$ -dimensional space  $S$ . The set of all vectors of the form  $A\mathbf{x}$ , where  $\mathbf{x} \in R$ , forms a vector space.<sup>10</sup> This space will be denoted by  $AR$ ; it is part of the space  $S$  or, as we shall say, is a *subspace* of  $S$ .

Together with the subspace  $AR$  of  $S$  we consider the set of all vectors  $\mathbf{x} \in R$  that satisfy the equation

$$A\mathbf{x} = \mathbf{o} \quad (38)$$

These vectors also form a subspace of  $R$ , which we shall denote by  $N_A$

<sup>10</sup> The set of vectors of the form  $A\mathbf{x}$  ( $\mathbf{x} \in R$ ) satisfies the postulates 1, 7, of § 1, because the sum of two such vectors and the product of such a vector by a number are also vectors of this form.

DEFINITION 9: If a linear operator  $A$  maps  $R$  into  $S$ , then the dimension  $r$  of the space  $AR$  is called the rank of  $A$ ,<sup>11</sup> and the dimension  $d$  of the space  $N_A$  consisting of all vectors  $x \in R$  that satisfy the condition (38) is called the defect, or nullity, of  $A$ .

Among all the equivalent rectangular matrices that describe a given operator  $A$  in distinct bases there occurs the canonical matrix  $I_r$  (see (37)). We denote the corresponding bases of  $R$  and  $S$  by  $e_1^*, e_2^*, \dots, e_n^*$  and  $g_1^*, g_2^*, \dots, g_m^*$ . Then

$$Ae_1^* = g_1^*, \dots, Ae_r^* = g_r^*, Ae_{r+1}^* = \dots = Ae_n^* = 0.$$

From the definition of  $AR$  and  $N_A$  it follows that the vectors  $g_1^*, g_2^*, \dots, g_r^*$  form a basis of  $AR$  and that the vectors  $e_{r+1}^*, e_{r+2}^*, \dots, e_n^*$  form a basis of  $N_A$ . Hence it follows that  $r$  is the rank of the operator  $A$  and that

$$d = n - r. \quad (39)$$

If  $A$  is an arbitrary matrix corresponding to  $A$ , then it is equivalent to  $I_r$  and therefore has the same rank  $r$ . Thus, the rank of an operator  $A$  coincides with the rank of the rectangular matrix  $A$

$$A = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{vmatrix}.$$

determined by  $A$  in arbitrary bases  $e_1, e_2, \dots, e_n \in R$  and  $g_1, g_2, \dots, g_m \in S$ .

The columns of  $A$  are formed by the coordinate vectors  $A_1e_1, \dots, A_ne_n$ . Since it follows from  $x = \sum_{i=1}^n x_i e_i$  that  $Ax = \sum_{i=1}^n x_i Ae_i$ , the rank of  $A$ , i.e., the dimension of  $RA$ , is equal to the maximal number of linearly independent vectors among  $Ae_1, Ae_2, \dots, Ae_n$ . Thus:

*The rank of a matrix coincides with the number of linearly independent columns of the matrix.*

Since under transposition the rows of a matrix become its columns and the rank remains unchanged:

<sup>11</sup> The dimension of the space  $AR$  never exceeds the dimension of  $R$ , so that  $r \leq n$ .

This follows from the fact that the equation  $x = \sum_{i=1}^n x_i e_i$  (where  $e_1, e_2, \dots, e_n$  is a basis of  $R$ ) implies the equation  $Ax = \sum_{i=1}^n x_i Ae_i$ .

The number of linearly independent rows of a matrix is also equal to the rank of the matrix.<sup>12</sup>

4. Let  $A$  and  $B$  be two linear operators and let  $C = AB$  be their product. Suppose that the operator  $B$  maps  $R$  into  $S$  and that the operator  $A$  maps  $S$  into  $T$ . Then the operator  $C$  maps  $R$  into  $T$ :

$$R \xrightarrow{B} S \xrightarrow{A} T, \quad R \xrightarrow{C} T.$$

We introduce the matrices  $A, B, C$  corresponding to  $A, B, C$  in some choice of bases in  $R, S$ , and  $T$ . Then the matrix equation  $C = AB$  will correspond to the operator equation  $C = AB$ .

We denote by  $r_A, r_B, r_C$  the ranks of the operators  $A, B, C$  or, what is the same, of the matrices  $A, B, C$ . These numbers determine the dimensions of the subspaces  $AS, BR, A(BR)$ . Since  $BR \subset S$ , we have  $A(BR) \subset AS$ .<sup>13</sup> Moreover, the dimension of  $A(BR)$  cannot exceed the dimension of  $BR$ .<sup>14</sup> Therefore

$$r_C \leq r_A, \quad r_C \leq r_B.$$

These inequalities were obtained in Chapter I, § 2 from the formula for the minors of a product of two matrices.

Let us regard  $A$  as an operator mapping  $BR$  into  $T$ . Then the rank of this operator is equal to the dimension of the space  $A(BR)$ , i.e., to  $r_C$ . Therefore, by applying (39) we obtain

$$Ax = o, \quad (40)$$

where  $d_1$  is the maximal number of linearly independent vectors of  $BR$  that satisfy the equation

$$r_C = r_B - d_1. \quad (41)$$

But all the solutions of this equation that belong to  $S$  form a subspace of dimension  $d$ , where

$$d = n - r_A \quad (42)$$

is the defect of the operator  $A$  mapping  $S$  into  $T$ . Since  $BR \subset S$ ,

$$d_1 \leq d. \quad (43)$$

From (41), (42), and (43) we find:

$$r_A + r_B - n \leq r_C.$$

<sup>12</sup> In § 1 we reached these conclusions on the basis of other arguments (see p. 54).

<sup>13</sup>  $R \subset S$  means that the set  $R$  forms part of the set  $S$ .

<sup>14</sup> See Footnote 11.

Thus we have obtained *Sylvester's inequality* for the rank of the product of two rectangular matrices  $A$  and  $B$  of dimensions  $m \times n$  and  $n \times q$ :

$$r_A + r_B - n \leq r_{AB} \leq \min(r_A, r_B). \quad (44)$$

### § 6. Linear Operators Mapping an $n$ -Dimensional Space into Itself

1. A linear operator mapping the  $n$ -dimensional vector space  $\mathbf{R}$  into itself (here  $\mathbf{R} \equiv \mathbf{S}$  and  $n = m$ ) will be referred to simply as a *linear operator in  $\mathbf{R}$* .

The sum of two linear operators in  $\mathbf{R}$  and the product of such an operator by a number are also linear operators in  $\mathbf{R}$ . Multiplication of two such operators is always feasible, and this product is also a linear operator in  $\mathbf{R}$ . Hence the linear operators in  $\mathbf{R}$  form a ring.<sup>15</sup> This ring has an identity operator, namely the operator  $\mathbf{E}$  for which

$$\mathbf{E}x = x \quad (x \in \mathbf{R}). \quad (45)$$

For every operator  $\mathbf{A}$  in  $\mathbf{B}$  we have

$$\mathbf{E}\mathbf{A} = \mathbf{A}\mathbf{E} = \mathbf{A}.$$

If  $\mathbf{A}$  is a linear operator in  $\mathbf{R}$ , then the powers  $\mathbf{A}^2 = \mathbf{A}\mathbf{A}$ ,  $\mathbf{A}^3 = \mathbf{A}\mathbf{A}\mathbf{A}$ , and in general  $\mathbf{A}^m = \underbrace{\mathbf{A}\mathbf{A} \cdots \mathbf{A}}_{m \text{ times}}$  have a meaning. We set  $\mathbf{A}^0 = \mathbf{E}$ . Then it is easy to see that for all non-negative integers  $p$  and  $q$  we have

$$\mathbf{A}^p \mathbf{A}^q = \mathbf{A}^{p+q}.$$

Let  $f(t) = \alpha_0 t^m + \alpha_1 t^{m-1} + \cdots + \alpha_{m-1} t + \alpha_m$  be a polynomial in a scalar argument  $t$  with coefficients in the field  $\mathbf{F}$ . Then we set:

$$f(\mathbf{A}) = \alpha_0 \mathbf{A}^m + \alpha_1 \mathbf{A}^{m-1} + \cdots + \alpha_{m-1} \mathbf{A} + \alpha_m \mathbf{E}. \quad (46)$$

Here  $f(\mathbf{A})g(\mathbf{A}) = g(\mathbf{A})f(\mathbf{A})$  for any two polynomials  $f(t)$  and  $g(t)$ .

Let

$$\mathbf{y} = \mathbf{A}x \quad (x, \mathbf{y} \in \mathbf{R}).$$

We denote by  $x_1, x_2, \dots, x_n$  the coordinates of the vector  $x$  in an arbitrary basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  and by  $y_1, y_2, \dots, y_n$  the coordinates of  $\mathbf{y}$  in the same basis. Then

$$y_i = \sum_{k=1}^n a_{ik} x_k \quad (i=1, 2, \dots, n). \quad (47)$$

<sup>15</sup> This ring is in fact an algebra. See Chapter I, p. 17.

In the basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  the linear operator  $\mathbf{A}$  corresponds to a square matrix  $A = \|a_{ik}\|_1^n$ .<sup>16</sup> We remind the reader (see § 2.) that in the  $k$ -th column of this matrix are to be found the coordinates of the vector  $\mathbf{A}\mathbf{e}_k$  ( $k = 1, 2, \dots, n$ ). Introducing the coordinate columns  $x = (x_1, x_2, \dots, x_n)$  and  $y = (y_1, y_2, \dots, y_n)$ , we can write the transformation (47) in matrix form

$$y = Ax. \quad (48)$$

The sum and product of two operators  $\mathbf{A}$  and  $\mathbf{B}$  correspond to the sum and product of the corresponding square matrices  $A = \|a_{ik}\|_1^n$  and  $B = \|b_{ik}\|_1^n$ . The product  $\alpha\mathbf{A}$  corresponds to the matrix  $\alpha A$ . The identity operator  $\mathbf{E}$  corresponds to the square unit matrix  $E = \|\delta_{ik}\|_1^n$ . Thus, the choice of a basis establishes an isomorphism between the ring of linear operators in  $\mathbf{R}$  and the ring of square matrices of order  $n$  with elements in  $\mathbf{F}$ . In this isomorphism the polynomial  $f(\mathbf{A})$  corresponds to the matrix  $f(A)$ .

Let us consider, apart from the basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ , another basis  $\mathbf{e}_1^*, \mathbf{e}_2^*, \dots, \mathbf{e}_n^*$  of  $\mathbf{R}$ . Then, in analogy with (48), we have

$$y^* = A^*x^*, \quad (49)$$

where  $x^*, y^*$  are the column matrices formed from the coordinates of the vectors  $\mathbf{x}, \mathbf{y}$  in the basis  $\mathbf{e}_1^*, \mathbf{e}_2^*, \dots, \mathbf{e}_n^*$  and  $A^* = \|a_{ik}^*\|_1^n$  is the square matrix corresponding to the operator  $\mathbf{A}$  in this basis. We rewrite in matrix form the formulas for the transformation of coordinates

$$x = Tx^*, \quad y = Ty^*. \quad (50)$$

Then from (48) and (50) we find:

$$y^* = T^{-1}ATx^*;$$

and a comparison with (49) gives:

$$A^* = T^{-1}AT. \quad (51)$$

Formula (51) is a special case of (31) on p. 61 (namely,  $P = T^{-1}$  and  $Q = T$ ).

DEFINITION 10: Two matrices  $A$  and  $B$  connected by the relation

$$B = T^{-1}AT, \quad (51')$$

where  $T$  is a non-singular matrix, are called similar.<sup>17</sup>

<sup>16</sup> See § 2 of this chapter. In this case the spaces  $\mathbf{R}$  and  $\mathbf{S}$  coincide; in the same way, the bases  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  and  $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n$  of these spaces are identified.

Thus, we have shown that *two matrices corresponding to one and the same linear operator in  $\mathbf{R}$  for distinct bases are similar* and the matrix  $T$  linking these matrices coincides with the matrix of the coordinate transformation in the transition from the first basis to the second (see (50)).

In other words, to a linear operator in  $\mathbf{R}$  there corresponds a whole class of similar matrices; they represent the given operator in various bases.

In studying properties of a linear operator in  $\mathbf{R}$ , we are at the same time studying the matrix properties that are common to the whole class of similar matrices, that is, that remain unchanged, or invariant, under transition from a given matrix to a similar one.

We note at once that two similar matrices always have the same determinant. For it follows from (51') that

$$B = |T|^{-1} |A| |T| = |A|. \quad (52)$$

The equation  $|B| = |A|$  is a necessary, but not a sufficient condition for the similarity of the matrices  $A$  and  $B$ .

In Chapter VI we shall establish a criterion for the similarity of two matrices, i.e., we shall give necessary and sufficient conditions for two square matrices of order  $n$  to be similar.

In accordance with (52) we may define the determinant  $|A|$  of a linear operator  $A$  in  $\mathbf{R}$  as the determinant of an arbitrary matrix corresponding to the given operator.

If  $|A| = 0$  ( $\neq 0$ ), then the operator  $A$  is called *singular* (*non-singular*). In accordance with this definition a singular (non-singular) operator corresponds to a singular (non-singular) matrix in any basis. For a singular operator:

- 1) There always exists a vector  $x \neq o$  such that  $Ax = o$ ;
- 2)  $AR$  is a proper part of  $\mathbf{R}$ .

For a non-singular operator:

- 1)  $Ax = o$  implies that  $x = o$ ;
- 2)  $AR \equiv \mathbf{R}$ , i.e., the vectors of the form  $Ax$  ( $x \in \mathbf{R}$ ) fill out the whole space  $\mathbf{R}$ .

In other words, a linear operator in  $\mathbf{R}$  is singular or non-singular depending on whether its defect is positive or zero.

---

<sup>17</sup> The matrix  $T$  can always be chosen such that its elements belong to the same basic number field  $\mathbf{F}$  as those of  $A$  and  $B$ . It is easy to verify the three properties of similar matrices:

*Reflexivity* (a matrix  $A$  is always similar to itself);  
*Symmetry* (if  $A$  is similar to  $B$ , then  $B$  is similar to  $A$ ); and  
*Transitivity* (if  $A$  is similar to  $B$ , and  $B$  to  $C$ , then  $A$  is similar to  $C$ ).





The equation (56) is an algebraic equation of degree  $n$  in  $\lambda$ . Its coefficients belong to the same number field  $\mathbf{F}$  as the elements of the matrix  $A = \|a_{ik}\|_1^n$ .

Equation (56) occurs in various problems of geometry, mechanics, astronomy, and physics and is known as the *characteristic equation* or the *secular equation*<sup>18</sup> of the matrix  $A = \|a_{ik}\|_1^n$  (the left-hand side is called the *characteristic polynomial*).

Thus, every characteristic value  $\lambda$  of a linear operator  $A$  is a root of the characteristic equation (56). And conversely, if a number  $\lambda$  is a root of (56), then for this value  $\lambda$  the system (55) and hence (54) has a non-zero solution  $x_1, x_2, \dots, x_n$ , i.e., to this number  $\lambda$  there corresponds a characteristic vector  $x = \sum x_i e_i$  of the operator  $A$ .

From what we have shown, it follows that every linear operator  $A$  in  $\mathbf{R}$  has not more than  $n$  distinct characteristic values.

If  $\mathbf{F}$  is the field of complex numbers, then every linear operator in  $\mathbf{R}$  always has at least one characteristic vector in  $\mathbf{R}$  corresponding to a characteristic value  $\lambda$ .<sup>19</sup> This follows from the fundamental theorem of algebra, according to which an algebraic equation (56) in the field of complex numbers always has at least one root.

Let us write (56) in explicit form

$$|A - \lambda E| \equiv (-\lambda)^n + S_1(-\lambda)^{n-1} + S_2(-\lambda)^{n-2} + \dots + S_{n-1}(-\lambda) + S_n = 0. \quad (57)$$

It is easy to see that here

$$S_1 = \sum_{i=1}^n a_{ii}, \quad S_2 = \sum_{1 \leq i < k \leq n} A \begin{pmatrix} i & k \\ i & k \end{pmatrix}, \quad \dots \quad (58)$$

and, in general,  $S_p$  is the sum of the principal minors of order  $p$  of the matrix  $A = \|a_{ik}\|_1^n$  ( $p = 1, 2, \dots, n$ ).<sup>20</sup> In particular,  $S_n = |A|$ .

We denote by  $\tilde{A}$  the matrix corresponding to the same operator  $A$  in another basis.  $\tilde{A}$  is similar to  $A$ :

<sup>18</sup> The name is due to the fact that this equation occurs in the study of secular perturbations of the planets.

<sup>19</sup> This proposition is valid even in the more general case in which  $\mathbf{F}$  is an arbitrary algebraically closed field, i.e., a field that contains the roots of all algebraic equations with coefficients in the field.

<sup>20</sup> The power  $(-\lambda)^{n-p}$  occurs only in those terms of the characteristic determinant (56) that contain precisely  $n-p$  of the diagonal elements, say,

$$a_{j_1 j_1} - \lambda, a_{j_2 j_2} - \lambda, \dots, a_{j_{n-p} j_{n-p}} - \lambda.$$

The product of these diagonal elements occurs in the expansion of the determinant (56)

$$\tilde{A} = T^{-1} A T.$$

Hence

$$\tilde{A} - \lambda E = T^{-1} (A - \lambda E) T$$

and therefore

$$|\tilde{A} - \lambda E| = |A - \lambda E|. \quad (59)$$

Thus, similar matrices  $A$  and  $\tilde{A}$  have the same characteristic polynomial. This polynomial is sometimes called the characteristic polynomial of the operator  $A$  and is denoted by  $|A - \lambda E|$ .

If  $x, y, z, \dots$  are linearly independent characteristic vectors of an operator  $A$  corresponding to one and the same characteristic  $\lambda$ , and  $\alpha, \beta, \gamma, \dots$  are arbitrary numbers of  $\mathbb{R}$ , then the vector  $\alpha x + \beta y + \gamma z + \dots$  is either equal to zero or is also a characteristic vector of  $A$  corresponding to the same  $\lambda$ .

For from

$$Ax = \lambda x, Ay = \lambda y, Az = \lambda z, \dots$$

it follows that

$$A(\alpha x + \beta y + \gamma z + \dots) = \lambda(\alpha x + \beta y + \gamma z + \dots).$$

In other words, linearly independent characteristic vectors corresponding to one and the same characteristic value  $\lambda$  form a basis of a 'characteristic' subspace each vector of which is a characteristic vector for the same  $\lambda$ . In particular, each characteristic vector generates a one-dimensional subspace, a 'characteristic' direction.

However, if characteristic vectors of a linear operator  $A$  correspond to distinct characteristic values, then a linear combination of these characteristic vectors is not, in general, a characteristic vector of  $A$ .

The significance of the characteristic vectors and characteristic numbers for the study of linear operators will be illustrated in the next section by the example of operators of simple structure.

with a factor in which the term free of  $\lambda$  is the principal minor

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix},$$

where  $i_1, i_2, \dots, i_p$  together with  $j_1, j_2, \dots, j_{n-p}$  forms a complete set of indices  $1, 2, \dots, n$ ; hence in the development of (56) we have

$$|A - \lambda E| = (a_{j_1 j_1} - \lambda) (a_{j_2 j_2} - \lambda) \dots (a_{j_{n-p} j_{n-p}} - \lambda) A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} + (*)$$

When we take all possible combinations  $j_1, j_2, \dots, j_{n-p}$  of  $n - p$  of the indices  $1, 2, \dots, n$ , we obtain for the coefficient  $S_p$  of  $(-\lambda)^{n-p}$  the sum of all principal minors of order  $p$  in  $A$ .

## § 8. Linear Operators of Simple Structure

1. We begin with the following lemma.

LEMMA: *Characteristic vectors belonging to pairwise distinct characteristic values are always linearly independent.*

*Proof.* Let

$$Ax_i = \lambda_i x_i \quad (x_i \neq 0; \lambda_i \neq \lambda_k \text{ for } i \neq k; i, k = 1, 2, \dots, m) \quad (60)$$

and

$$\sum_{i=1}^m c_i x_i = 0. \quad (61)$$

Applying the operator  $A$  to both sides we obtain:

$$\sum_{i=1}^m c_i \lambda_i x_i = 0. \quad (62)$$

We multiply both sides of (61) by  $\lambda_1$  and subtract (61) from (62) term by term. Then we obtain

$$\sum_{i=2}^m c_i (\lambda_i - \lambda_1) x_i = 0. \quad (63)$$

We can say that (63) is obtained from (61) by termwise application of the operator  $A - \lambda_1 E$ . If we apply the operators  $A - \lambda_2 E, \dots, A - \lambda_{m-1} E$  to (63) term by term, we are led to the following equation:

$$c_m (\lambda_m - \lambda_{m-1}) (\lambda_m - \lambda_{m-2}) \cdots (\lambda_m - \lambda_1) x_m = 0,$$

so that  $c_m = 0$ . Since any of the summands in (61) can be put last, we have in (61)

$$c_1 = c_2 = \dots = c_m = 0,$$

i.e., there is no linear dependence among the vectors  $x_1, x_2, \dots, x_m$ . This proves the lemma.

If the characteristic equation of an operator has  $n$  distinct roots and these roots belong to  $\mathbb{F}$ , then by the lemma the characteristic vectors belonging to these roots are linearly independent.

DEFINITION 11: A linear operator  $A$  in  $\mathbf{R}$  is said to be an operator of simple structure if  $A$  has  $n$  linearly independent characteristic vectors in  $\mathbf{R}$ , where  $n$  is the dimension of  $\mathbf{R}$ .

Thus, a linear operator in  $\mathbf{R}$  has simple structure if all the roots of the characteristic equation are distinct and belong to  $\mathbb{F}$ . However, these condi-

tions are not necessary. There exist linear operators of simple structure whose characteristic polynomial has multiple roots.

Let us consider an arbitrary linear operator  $A$  of simple structure. We denote by  $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n$  a basis of  $\mathbf{R}$  consisting of characteristic vectors of the operator, i.e.,

$$A\mathbf{g}_k = \lambda_k \mathbf{g}_k \quad (k=1, 2, \dots, n).$$

If

$$\mathbf{x} = \sum_{k=1}^n x_k \mathbf{g}_k,$$

then

$$A\mathbf{x} = \sum_{k=1}^n x_k A\mathbf{g}_k = \sum_{k=1}^n \lambda_k x_k \mathbf{g}_k.$$

The effect of the operator  $A$  of simple structure on the vector  $\mathbf{x} = \sum_{k=1}^n x_k \mathbf{g}_k$  may be put into words as follows:

*In the  $n$ -dimensional space  $\mathbf{R}$  there exist  $n$  linearly independent 'directions' along which the operator  $A$  of simple structure realizes a 'dilatation' with coefficients  $\lambda_1, \lambda_2, \dots, \lambda_n$ . An arbitrary vector  $\mathbf{x}$  may be decomposed into components along these characteristic directions. These components are subject to the corresponding 'dilatations' and their sum then gives the vector  $A\mathbf{x}$ .*

It is easy to see that to the operator  $A$  in a 'characteristic' basis  $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n$  there corresponds the diagonal matrix

$$\tilde{A} = \|\lambda_i \delta_{ik}\|_1^n.$$

If we denote by  $A$  the matrix corresponding to  $A$  in an arbitrary basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ , then

$$A = T \|\lambda_i \delta_{ik}\|_1^n T^{-1}. \quad (64)$$

A matrix that is similar (p. 68) to a diagonal matrix is called a *matrix of simple structure*. Thus, to an operator of simple structure there corresponds in any basis a matrix of simple structure, and vice versa.

**2.** The matrix  $T$  in (64) realizes the transition from the basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  to the basis  $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n$ . The  $k$ -th column of  $T$  contains the coordinates of a characteristic vector  $\mathbf{g}_k$  (with respect to  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ ) that corresponds to the characteristic value  $\lambda_k$  of  $A$  ( $k=1, 2, \dots, n$ ). The matrix  $T$  is called the *fundamental matrix* for  $A$ .

We rewrite (64) as follows:

$$A = TLT^{-1} \quad (L = \{\lambda_1, \lambda_2, \dots, \lambda_n\}). \quad (64')$$

On going over to the  $p$ -th compound matrix ( $1 \leq p \leq n$ ), we obtain (see Chapter I, § 4):

$$\mathfrak{A}_p = \mathfrak{T}_p \mathfrak{L}_p \mathfrak{T}_p^{-1}. \quad (65)$$

$\mathfrak{L}_p$  is a diagonal matrix of order  $N$  ( $N = \binom{n}{p}$ ) along whose main diagonal are all the possible products of  $\lambda_1, \lambda_2, \dots, \lambda_n$  taken  $p$  at a time. A comparison of (65) with (64') yields the following theorem:

**THEOREM 3:** *If a matrix  $A = \|a_{ik}\|_1^n$  has simple structure, then for every  $p \leq n$  the compound matrix  $\mathfrak{A}_p$  also has simple structure; moreover, the characteristic values of  $\mathfrak{A}_p$  are all the possible products  $\lambda_{i_1} \lambda_{i_2} \dots \lambda_{i_p}$  ( $1 \leq i_1 < i_2 < \dots < i_p \leq n$ ) of  $p$  of the characteristic values  $\lambda_1, \lambda_2, \dots, \lambda_n$  of  $A$ , and the fundamental matrix of  $\mathfrak{A}_p$  is the compound  $\mathfrak{T}_p$  of the fundamental matrix  $T$  of  $A$ .*

**COROLLARY:** *If a characteristic value  $\lambda_k$  of a matrix of simple structure  $A = \|a_{ik}\|_1^n$  corresponds to a characteristic vector with the coordinates  $t_{1k}, t_{2k}, \dots, t_{nk}$  ( $k = 1, 2, \dots, n$ ) and if  $T = \|t_{ik}\|_1^n$ , then the characteristic value  $\lambda_{k_1} \lambda_{k_2} \dots \lambda_{k_p}$  ( $1 \leq k_1 < k_2 < \dots < k_p \leq n$ ) of  $\mathfrak{A}_p$  corresponds to the characteristic vector with coordinates*

$$T \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \quad (1 \leq i_1 < i_2 < \dots < i_p \leq n). \quad (66)$$

An arbitrary matrix  $A = \|a_{ik}\|_1^n$  may be represented in the form of a sequence of matrices  $A_m$  ( $m \rightarrow \infty$ ) each of which does not have multiple characteristic values and, therefore, has simple structure. The characteristic values  $\lambda_1^{(m)}, \lambda_2^{(m)}, \dots, \lambda_n^{(m)}$  of the matrix  $A_m$  converge for  $m \rightarrow \infty$  to the characteristic values  $\lambda_1, \lambda_2, \dots, \lambda_n$  of  $A$ ,

$$\lim_{m \rightarrow \infty} \lambda_k^{(m)} = \lambda_k \quad (k = 1, 2, \dots, n).$$

Hence

$$\lim_{m \rightarrow \infty} \lambda_{k_1}^{(m)} \lambda_{k_2}^{(m)} \dots \lambda_{k_p}^{(m)} = \lambda_{k_1} \lambda_{k_2} \dots \lambda_{k_p} \quad (1 \leq k_1 < k_2 < \dots < k_p \leq n).$$

Moreover, since  $\lim_{m \rightarrow \infty} \mathfrak{A}_{(m)p} = \mathfrak{A}_p$ , we deduce from Theorem 3:

**THEOREM 4 (Kronecker):** *If  $\lambda_1, \lambda_2, \dots, \lambda_n$  is a complete system of characteristic values of an arbitrary matrix  $A$ , then a complete system of characteristic values of the compound matrix  $\mathfrak{A}_p$  consists of all possible products of the numbers  $\lambda_1, \lambda_2, \dots, \lambda_n$  taken  $p$  at a time ( $p = 1, 2, \dots, n$ ).*

In the present section we have investigated operators and matrices of simple structure. The study of the structure of operators and matrices of general type will be resumed in Chapters VI and VII.

## CHAPTER IV

### THE CHARACTERISTIC POLYNOMIAL AND THE MINIMAL POLYNOMIAL OF A MATRIX

Two polynomials are associated with every square matrix: the characteristic polynomial and the minimal polynomial. These polynomials play an important role in various problems of the theory of matrices. For example, the concept of a function of a matrix, which we shall introduce in the next chapter, will be based entirely on the concept of the minimal polynomial. In the present chapter, the properties of the characteristic polynomial and the minimal polynomial are studied. A prerequisite to this investigation is some basic information about polynomials with matrix coefficients and operations on them:

#### § 1. Addition and Multiplication of Matrix Polynomials

1. We consider a square *polynomial* matrix  $A(\lambda)$ , i.e., a square matrix whose elements are polynomials in  $\lambda$  (with coefficients in the given number field  $\mathbb{F}$ ):

$$A(\lambda) = \|a_{ik}(\lambda)\|_1^n = \|a_{ik}^{(0)}\lambda^m + a_{ik}^{(1)}\lambda^{m-1} + \dots + a_{ik}^{(m)}\|_1^n. \quad (1)$$

The matrix  $A(\lambda)$  can be represented in the form of a polynomial with matrix coefficients arranged with respect to the powers of  $\lambda$ :

$$A(\lambda) = A_0\lambda^m + A_1\lambda^{m-1} + \dots + A_m, \quad (2)$$

where

$$A_j = \|a_{ik}^{(j)}\|_1^n \quad (j = 0, 1, \dots, m). \quad (3)$$

The number  $m$  is called the *degree* of the polynomial, provided  $A_0 \neq 0$ . The number  $n$  is called the *order* of the polynomial. The polynomial (1) is called *regular* if  $|A_0| \neq 0$ .

A polynomial with matrix coefficients will sometimes be called a *matrix polynomial*. In contrast to a matrix polynomial an ordinary polynomial with scalar coefficients will be called a *scalar polynomial*.

We shall now consider the fundamental operations on matrix polynomials. Let two matrix polynomials  $A(\lambda)$  and  $B(\lambda)$  of the same order be given. We denote by  $m$  the larger of their degrees. These polynomials can be written in the form

$$\begin{aligned} A(\lambda) &= A_0\lambda^m + A_1\lambda^{m-1} + \dots + A_m, \\ B(\lambda) &= B_0\lambda^m + B_1\lambda^{m-1} + \dots + B_m. \end{aligned}$$

Then

$$A(\lambda) \pm B(\lambda) = (A_0 \pm B_0)\lambda^m + (A_1 \pm B_1)\lambda^{m-1} + \dots + (A_m \pm B_m),$$

i.e.: *The sum (difference) of two matrix polynomials of the same order can be represented in the form of a polynomial whose degree does not exceed the larger of the degrees of the given polynomials.*

Let  $A(\lambda)$  and  $B(\lambda)$  be two matrix polynomials of the same order  $n$  and of respective degrees  $m$  and  $p$ :

$$\begin{aligned} A(\lambda) &= A_0\lambda^m + A_1\lambda^{m-1} + \dots + A_m & (A_0 \neq O), \\ B(\lambda) &= B_0\lambda^p + B_1\lambda^{p-1} + \dots + B_p & (B_0 \neq O). \end{aligned}$$

Then

$$A(\lambda)B(\lambda) = A_0B_0\lambda^{m+p} + (A_0B_1 + A_1B_0)\lambda^{m+p-1} + \dots + A_mB_p. \quad (4)$$

If we multiply  $B(\lambda)$  by  $A(\lambda)$  (i.e., interchange the order of the factors), then we obtain, in general, a different polynomial.

**2.** The multiplication of matrix polynomials has a specific property. In contrast to the product of scalar polynomials, the product (4) of matrix polynomials may have a degree less than  $m+p$ , i.e., less than the sum of the degrees of the factors. For, in (4) the product  $A_0B_0$  may be the null matrix even though  $A_0 \neq O$ ,  $B_0 \neq O$ . However, if at least one of the matrices  $A_0$  and  $B_0$  is non-singular, then it follows from  $A_0 \neq O$  and  $B_0 \neq O$  that  $A_0B_0 \neq O$ . Thus: *The product of two matrix polynomials is a polynomial whose degree is less than or equal to the sum of the degrees of the factors. If at least one of the two factors is regular, then the degree of the product is always equal to the sum of the degrees of the factors.*

## § 2. Right and Left Division of Matrix Polynomials

**1.** Let  $A(\lambda)$  and  $B(\lambda)$  be two matrix polynomials of the same order  $n$ , and let  $B(\lambda)$  be regular:

$$\begin{aligned} A(\lambda) &= A_0\lambda^m + A_1\lambda^{m-1} + \dots + A_m & (A_0 \neq O), \\ B(\lambda) &= B_0\lambda^p + B_1\lambda^{p-1} + \dots + B_p & (|B_0| \neq 0). \end{aligned}$$



We shall say that the matrix polynomials  $Q(\lambda)$  and  $R(\lambda)$  are the *right quotient* and the *right remainder*, respectively, of  $A(\lambda)$  on division by  $B(\lambda)$  if

$$A(\lambda) = Q(\lambda) B(\lambda) + R(\lambda) \quad (5)$$

and if the degree of  $R(\lambda)$  is less than that of  $B(\lambda)$ .

Similarly, we shall call the polynomials  $\hat{Q}(\lambda)$  and  $\hat{R}(\lambda)$  the *left quotient* and the *left remainder* of  $A(\lambda)$  on division by  $B(\lambda)$  if

$$A(\lambda) = B(\lambda) \hat{Q}(\lambda) + \hat{R}(\lambda) \quad (6)$$

and if the degree of  $\hat{R}(\lambda)$  is less than that of  $B(\lambda)$ .

The reader should note that in the 'right' division (i.e., when the right quotient and the right remainder are to be found) in (5) the quotient  $Q(\lambda)$  is multiplied by the 'divisor'  $B(\lambda)$  on the *right*, and in the 'left' division in (6) the quotient  $\hat{Q}(\lambda)$  is multiplied by the divisor  $B(\lambda)$  on the *left*. The polynomials  $\hat{Q}(\lambda)$  and  $\hat{R}(\lambda)$  do not, in general, coincide with  $Q(\lambda)$  and  $R(\lambda)$ .

**2.** We shall now show that *both right and left division of matrix polynomials of the same order are always possible and unique, provided the divisor is a regular polynomial*.

Let us consider the right division of  $A(\lambda)$  by  $B(\lambda)$ . If  $m < p$ , we can set  $Q(\lambda) = 0$ ,  $R(\lambda) = A(\lambda)$ . If  $m \geq p$ , we apply the usual scheme for the division of a polynomial by a polynomial in order to find the quotient  $Q(\lambda)$  and the remainder  $R(\lambda)$ . We 'divide' the highest term of the dividend  $A_0 \lambda^m$  by the highest term of the divisor  $B_0 \lambda^p$ . We obtain the highest term  $A_0 B_0^{-1} \lambda^{m-p}$  of the required quotient. We multiply this term on the right by the divisor  $B(\lambda)$  and subtract the product so obtained from  $A(\lambda)$ . Thus we find the 'first remainder'  $A^{(1)}(\lambda)$ :

$$A(\lambda) = A_0 B_0^{-1} \lambda^{m-p} B(\lambda) + A^{(1)}(\lambda). \quad (7)$$

The degree  $m^{(1)}$  of  $A^{(1)}(\lambda)$  is less than  $m$ :

$$A^{(1)}(\lambda) = A_0^{(1)} \lambda^{m^{(1)}} + \dots \quad (A_0^{(1)} \neq 0, \quad m^{(1)} < m). \quad (8)$$

If  $m^{(1)} \geq p$ , then we repeat the process and obtain:

$$\left. \begin{aligned} A^{(1)}(\lambda) &= A_0^{(1)} B_0^{-1} \lambda^{m^{(1)}-p} B(\lambda) + A^{(2)}(\lambda), \\ A^{(2)}(\lambda) &= A_0^{(2)} \lambda^{m^{(2)}} + \dots \quad (m^{(2)} < m^{(1)}), \end{aligned} \right\} \quad (9)$$

etc.

Since the degrees of  $A(\lambda)$ ,  $A^{(1)}(\lambda)$ ,  $A^{(2)}(\lambda)$ , ... decrease, at some stage we arrive at a remainder  $R(\lambda)$  whose degree is less than  $p$ . Then it follows from (7) and (9) that

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda),$$

where

$$Q(\lambda) = A_0 B_0^{-1} \lambda^{m-p} + A_0^{(1)} B_0^{-1} \lambda^{m(1)-p} + \dots \quad (10)$$

We shall now prove the *uniqueness* of the right division. Suppose we have simultaneously

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda) \quad (11)$$

and

$$A(\lambda) = Q^*(\lambda)B(\lambda) + R^*(\lambda), \quad (12)$$

where the degrees of  $R(\lambda)$  and  $R^*(\lambda)$  are less than that of  $B(\lambda)$ , i.e., less than  $p$ . Subtracting (11) from (12) term by term we obtain

$$[Q(\lambda) - Q^*(\lambda)]B(\lambda) = R^*(\lambda) - R(\lambda). \quad (13)$$

If we had  $Q(\lambda) - Q^*(\lambda) \neq 0$ , then the degree on the left-hand side of (13) would be the sum of the degrees of  $B(\lambda)$  and  $Q(\lambda) - Q^*(\lambda)$ , because  $|B_0| \neq 0$ , and would therefore be at least equal to  $p$ . This is impossible, since the degree of the polynomial on the right-hand side of (13) is less than  $p$ . Thus,  $Q(\lambda) - Q^*(\lambda) \equiv 0$ , and then it follows from (13) that  $R(\lambda) - R^*(\lambda) \equiv 0$ , i.e.,

$$Q(\lambda) = Q^*(\lambda), \quad R(\lambda) = R^*(\lambda).$$

The existence and uniqueness of the left quotient and left remainder is established similarly.<sup>1</sup>

<sup>1</sup> Note that the possibility and uniqueness of the left division of  $A(\lambda)$  by  $B(\lambda)$  follows from that of the right division of the transposed matrices  $A^T(\lambda)$  and  $B^T(\lambda)$ . (The regularity of  $B(\lambda)$  implies that of  $B^T(\lambda)$ .) For from

$$A^T(\lambda) = Q_1(\lambda)B^T(\lambda) + R_1(\lambda)$$

it follows (see Chapter I, p. 19) that

$$A(\lambda) = B(\lambda)Q_1^T(\lambda) + R_1^T(\lambda). \quad (6')$$

By the same reasoning, the left division of  $A(\lambda)$  by  $B(\lambda)$  is unique; for if it were not, then the right division of  $A^T(\lambda)$  by  $B^T(\lambda)$  would not be unique.

Comparison of (6) and (6') gives

$$\hat{Q}(\lambda) = Q_1^T(\lambda), \quad \hat{R}(\lambda) = R_1^T(\lambda).$$

*Example.*

$$A(\lambda) = \left\| \begin{array}{cc} \lambda^3 + \lambda & 2\lambda^3 + \lambda^2 \\ -\lambda^3 - 2\lambda^2 + 1 & 3\lambda^3 + \lambda \end{array} \right\|$$

$$= \overbrace{\left\| \begin{array}{cc} 1 & 2 \\ -1 & 3 \end{array} \right\|}^{A_0} \lambda^3 + \left\| \begin{array}{cc} 0 & 1 \\ -2 & 0 \end{array} \right\| \lambda^2 + \left\| \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right\| \lambda + \left\| \begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array} \right\|,$$

$$B(\lambda) = \left\| \begin{array}{cc} 2\lambda^3 + 3 & -\lambda^3 + 1 \\ -\lambda^3 - 1 & \lambda^3 + 2 \end{array} \right\| = \overbrace{\left\| \begin{array}{cc} 2 & -1 \\ -1 & 1 \end{array} \right\|}^{B_0} \lambda^3 + \left\| \begin{array}{cc} 3 & 1 \\ -1 & 2 \end{array} \right\|,$$

$$|B_0| = 1, B_0^{-1} = \left\| \begin{array}{cc} 1 & 1 \\ 1 & 2 \end{array} \right\|, A_0 B_0^{-1} = \left\| \begin{array}{cc} 3 & 5 \\ 2 & 5 \end{array} \right\|, A_0 B_0^{-1} B(\lambda) = \left\| \begin{array}{cc} \lambda^2 + 4 & 2\lambda^3 + 13 \\ -\lambda^3 + 1 & 3\lambda^3 + 12 \end{array} \right\|,$$

$$A^{(1)}(\lambda) = \left\| \begin{array}{cc} \lambda^3 + \lambda & 2\lambda^3 + \lambda^2 \\ -\lambda^3 - 2\lambda^2 + 1 & 3\lambda^3 + \lambda \end{array} \right\| - \left\| \begin{array}{cc} \lambda^3 + 4\lambda & 2\lambda^3 + 13\lambda \\ -\lambda^3 + \lambda & 3\lambda^3 + 12\lambda \end{array} \right\| \\ = \left\| \begin{array}{cc} -3\lambda & \lambda^2 - 13\lambda \\ -2\lambda^3 - \lambda + 1 & -11\lambda \end{array} \right\|,$$

$$A^{(1)}(\lambda) = \left\| \begin{array}{cc} 0 & 1 \\ -2 & 0 \end{array} \right\| \lambda^2 + \left\| \begin{array}{cc} -3 & -13 \\ -1 & -11 \end{array} \right\| \lambda + \left\| \begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array} \right\|,$$

$$A_0^{(1)} B_0^{-1} = \left\| \begin{array}{cc} 0 & 1 \\ -2 & 0 \end{array} \right\| \cdot \left\| \begin{array}{cc} 1 & 1 \\ 1 & 2 \end{array} \right\| = \left\| \begin{array}{cc} 1 & 2 \\ -2 & -2 \end{array} \right\|,$$

$$A_0^{(1)} B_0^{-1} B(\lambda) = \left\| \begin{array}{cc} 1 & 2 \\ -2 & -2 \end{array} \right\| \cdot \left\| \begin{array}{cc} 2\lambda^3 + 3 & -\lambda^3 + 1 \\ -\lambda^3 - 1 & \lambda^3 + 2 \end{array} \right\| = \left\| \begin{array}{cc} 1 & \lambda^3 + 5 \\ -2\lambda^3 - 4 & -6 \end{array} \right\|,$$

$$R(\lambda) = A^{(1)}(\lambda) - A_0^{(1)} B_0^{-1} B(\lambda)$$

$$= \left\| \begin{array}{cc} -3\lambda & \lambda^2 - 13\lambda \\ -2\lambda^3 - \lambda + 1 & -11\lambda \end{array} \right\| - \left\| \begin{array}{cc} 1 & \lambda^3 + 5 \\ -2\lambda^3 - 4 & -6 \end{array} \right\| = \left\| \begin{array}{cc} -3\lambda - 1 & -13\lambda - 5 \\ -\lambda + 5 & -11\lambda + 6 \end{array} \right\|,$$

$$Q(\lambda) = A_0 B_0^{-1} \lambda + A_0^{(1)} B_0^{-1} = \left\| \begin{array}{cc} 3 & 5 \\ 2 & 5 \end{array} \right\| \lambda + \left\| \begin{array}{cc} 1 & 2 \\ -2 & -2 \end{array} \right\| = \left\| \begin{array}{cc} 3\lambda + 1 & 5\lambda + 2 \\ 2\lambda - 2 & 5\lambda - 2 \end{array} \right\|.$$

As an exercise, the reader should verify that

$$A(\lambda) = Q(\lambda)B(\lambda) + R(\lambda).$$

### § 3. The Generalized Bézout Theorem

1. We consider an arbitrary matrix polynomial of order  $n$

$$F(\lambda) = F_0 \lambda^m + F_1 \lambda^{m-1} + \dots + F_m \quad (F_0 \neq O). \quad (14)$$

This polynomial can also be written as follows:

$$F(\lambda) = \lambda^m F_0 + \lambda^{m-1} F_1 + \dots + F_m. \quad (15)$$

For a scalar  $\lambda$ , both ways of writing give the same result. However, if we substitute for the scalar argument  $\lambda$  a square matrix  $A$  of order  $n$ , then the results of the substitution in (14) and (15) will, in general, be distinct, since the powers of  $A$  need not be permutable with the matrix coefficients  $F_0, F_1, \dots, F_m$ .

We set

$$F(A) = F_0 A^m + F_1 A^{m-1} + \dots + F_m \quad (16)$$

and

$$\widehat{F}(A) = A^m F_0 + A^{m-1} F_1 + \dots + F_m, \quad (17)$$

and call  $F(A)$  the *right value* and  $\widehat{F}(A)$  the *left value* of  $F(\lambda)$  on substitution of  $A$  for  $\lambda$ .<sup>2</sup>

We divide  $F(\lambda)$  by the binomial  $\lambda E - A$ . In this case the right remainder  $R(\lambda)$  and left remainder  $\widehat{R}(\lambda)$  will not depend on  $\lambda$ . To determine the right remainder we use the usual division scheme:

$$\begin{aligned} F(\lambda) &= F_0 \lambda^m + F_1 \lambda^{m-1} + \dots + F_m \\ &= F_0 \lambda^{m-1} (\lambda E - A) + (F_0 A + F_1) \lambda^{m-1} + F_2 \lambda^{m-2} + \dots \\ &= [F_0 \lambda^{m-1} + (F_0 A + F_1) \lambda^{m-2}] (\lambda E - A) + (F_0 A^2 + F_1 A + F_2) \lambda^{m-2} + F_3 \lambda^{m-3} + \dots \\ &= [F_0 \lambda^{m-1} + (F_0 A + F_1) \lambda^{m-2} + \dots \\ &\quad + F_0 A^{m-1} + F_1 A^{m-2} + \dots + F_{m-1}] (\lambda E - A) \\ &\quad + F_0 A^m + F_1 A^{m-1} + \dots + F_m. \end{aligned}$$

Thus we have found that

$$R = F_0 A^m + F_1 A^{m-1} + \dots + F_m = F(A). \quad (18)$$

Similarly

$$\widehat{R} = \widehat{F}(A). \quad (19)$$

This proves

**THEOREM 1** (The Generalized Bézout Theorem): *When the matrix polynomial  $F(\lambda)$  is divided on the right by the binomial  $\lambda E - A$ , the remainder is  $F(A)$ ; when it is divided on the left, the remainder is  $\widehat{F}(A)$ .*

---

<sup>2</sup> In the 'right' value  $F(A)$  the powers of  $A$  are at the right of the coefficients; in the 'left' value  $\widehat{F}(A)$ , at the left.

2. From this theorem it follows that:

A polynomial  $F(\lambda)$  is divisible by the binomial  $\lambda E - A$  on the right (left) without remainder if and only if  $F(A) = O$  ( $\hat{F}(A) = O$ ).

*Example.* Let  $A = \|a_{ik}\|_1^n$  and let  $f(\lambda)$  be a polynomial in  $\lambda$ . Then

$$F(\lambda) = f(\lambda)E - f(A)$$

is divisible by  $\lambda E - A$  (both on the right and on the left) without remainder. This follows immediately from the generalized Bézout Theorem, because in this case  $F(A) = \hat{F}(A) = O$ .

#### § 4. The Characteristic Polynomial of a Matrix. The Adjoint Matrix

1. We consider a matrix  $A = \|a_{ik}\|_1^n$ . The characteristic matrix of  $A$  is  $\lambda E - A$ . The determinant of the characteristic matrix

$$\Delta(\lambda) = |\lambda E - A| = |\lambda \delta_{ik} - a_{ik}|_1^n,$$

is a scalar polynomial in  $\lambda$  and is called the *characteristic polynomial* of  $A$  (see Chapter III, § 7).<sup>3</sup>

The matrix  $B(\lambda) = \|b_{ik}(\lambda)\|_1^n$ , where  $b_{ik}(\lambda)$  is the algebraic complement of the element  $\lambda \delta_{ik} - a_{ik}$  in the determinant  $\Delta(\lambda)$  is called the *adjoint matrix* of  $A$ .

By way of example, for the matrix

$$A = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$

we have:

$$\lambda E - A = \begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & \lambda - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda - a_{33} \end{vmatrix},$$

$$\Delta(\lambda) = |\lambda E - A| = \lambda^3 - (a_{11} + a_{22} + a_{33})\lambda^2 + \dots,$$

$$B(\lambda) = \begin{vmatrix} \lambda^2 - (a_{22} + a_{33})\lambda + a_{22}a_{33} - a_{23}a_{32} & * & * \\ a_{21}\lambda + a_{23}a_{31} - a_{21}a_{33} & * & * \\ a_{31}\lambda + a_{21}a_{32} - a_{22}a_{31} & * & * \end{vmatrix}.$$

<sup>3</sup> This polynomial differs by the factor  $(-1)^n$  from the polynomial  $\Delta(\lambda)$  introduced in Chapter III, § 7.

These definitions imply the following identities in  $\lambda$ :

$$(\lambda E - A) B(\lambda) = \Delta(\lambda) E, \quad (20)$$

$$B(\lambda) (\lambda E - A) = \Delta(\lambda) E. \quad (20')$$

The right-hand sides of these equations can be regarded as polynomials with matrix coefficients (each of these coefficients is the product of a scalar and the unit matrix  $E$ ). The polynomial matrix  $B(\lambda)$  can also be represented in the form of a polynomial arranged with respect to the powers of  $\lambda$ . Equations (20) and (20') show that  $\Delta(\lambda)E$  is divisible on the right and on the left by  $\lambda E - A$  without remainder. By the Generalized Bézout Theorem, this is only possible when the remainder  $\Delta(A)E = \Delta(A)$  is the null matrix. Thus we have proved:

**THEOREM 2 (Hamilton-Cayley):** *Every square matrix  $A$  satisfies its characteristic equation, i.e.*

$$\Delta(A) = O. \quad (21)$$

*Example.*

$$A = \begin{vmatrix} 2 & 1 \\ -1 & 3 \end{vmatrix},$$

$$\Delta(\lambda) = \begin{vmatrix} \lambda - 2 & -1 \\ 1 & \lambda - 3 \end{vmatrix} = \lambda^2 - 5\lambda + 7,$$

$$\Delta(A) = A^2 - 5A + 7E = \begin{vmatrix} 3 & 5 \\ -5 & 8 \end{vmatrix} - 5 \begin{vmatrix} 2 & 1 \\ -1 & 3 \end{vmatrix} + 7 \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix} = O.$$

**2.** We denote by  $\lambda_1, \lambda_2, \dots, \lambda_n$  all the characteristic values of  $A$ , i.e., all the roots of the characteristic polynomial  $\Delta(\lambda)$  (each  $\lambda_i$  is repeated as often as its multiplicity as a root of  $\Delta(\lambda)$  requires). Then

$$\Delta(\lambda) = |\lambda E - A| = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n). \quad (22)$$

Let  $g(\mu)$  be an arbitrary scalar polynomial. We wish to find the characteristic values of  $g(A)$ . For this purpose we split  $g(\mu)$  into linear factors

$$g(\mu) = a_0(\mu - \mu_1)(\mu - \mu_2) \cdots (\mu - \mu_t). \quad (23)$$

On both sides of this identity we substitute the matrix  $A$  for  $\mu$ :

$$g(A) = a_0(A - \mu_1 E)(A - \mu_2 E) \cdots (A - \mu_t E). \quad (24)$$

Passing to determinants on both sides of (24) and using (22) and (23) we find

$$\begin{aligned}
|g(A)| &= a_0^n |A - \mu_1 E| |A - \mu_2 E| \cdots |A - \mu_l E| \\
&= (-1)^{ni} a_0^n \Delta(\mu_1) \Delta(\mu_2) \cdots \Delta(\mu_l) \\
&= (-1)^{ni} a_0^n \prod_{i=1}^l \prod_{k=1}^{n_i} (\mu_i - \lambda_k) = g(\lambda_1) g(\lambda_2) \cdots g(\lambda_n).
\end{aligned}$$

If in the equation

$$|g(A)| = g(\lambda_1) g(\lambda_2) \cdots g(\lambda_n) \quad (25)$$

we replace the polynomial  $g(\mu)$  by  $\lambda - g(\mu)$ , where  $\lambda$  is some parameter, we find:

$$|\lambda E - g(A)| = [\lambda - g(\lambda_1)] [\lambda - g(\lambda_2)] \cdots [\lambda - g(\lambda_n)]. \quad (26)$$

This leads to the following theorem.

**THEOREM 3:** *If  $\lambda_1, \lambda_2, \dots, \lambda_n$  are all the characteristic values (with the proper multiplicities) of a matrix  $A$  and if  $g(\mu)$  is a scalar polynomial, then  $g(\lambda_1), g(\lambda_2), \dots, g(\lambda_n)$  are the characteristic values of  $g(A)$ .*

In particular, if  $A$  has the characteristic values  $\lambda_1, \lambda_2, \dots, \lambda_n$ , then  $A^k$  has the characteristic values  $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$  ( $k = 0, 1, 2, \dots$ ).

**3.** We shall now derive an effective formula expressing the adjoint matrix  $B(\lambda)$  in terms of the characteristic polynomial  $\Delta(\lambda)$ .

Let

$$\Delta(\lambda) = \lambda^n - p_1 \lambda^{n-1} - p_2 \lambda^{n-2} - \cdots - p_n. \quad (27)$$

The difference  $\Delta(\lambda) - \Delta(\mu)$  is divisible by  $\lambda - \mu$  without remainder. Therefore

$$\delta(\lambda, \mu) = \frac{\Delta(\lambda) - \Delta(\mu)}{\lambda - \mu} = \lambda^{n-1} + (\mu - p_1) \lambda^{n-2} + (\mu^2 - p_1 \mu - p_2) \lambda^{n-3} + \cdots \quad (28)$$

is a polynomial in  $\lambda$  and  $\mu$ .

The identity

$$\Delta(\lambda) - \Delta(\mu) = \delta(\lambda, \mu) (\lambda - \mu) \quad (29)$$

will still hold if we replace  $\lambda$  and  $\mu$  by the permutable matrices  $\lambda E$  and  $A$ . Since by the Hamilton-Cayley Theorem  $\Delta(A) = O$ ,

$$\Delta(\lambda) E = \delta(\lambda E, A) (\lambda E - A). \quad (30)$$

Comparing (20') with (30), we obtain by virtue of the uniqueness of the quotient the required formula

$$B(\lambda) = \delta(\lambda E, A). \quad (31)$$

Hence by (28)

$$B(\lambda) = E\lambda^{n-1} + B_1\lambda^{n-2} + B_2\lambda^{n-3} + \dots + B_{n-1}, \quad (32)$$

where

$$B_1 = A - p_1 E, \quad B_2 = A^2 - p_1 A - p_2 E, \quad \dots$$

and, in general,

$$B_k = A^k - p_1 A^{k-1} - p_2 A^{k-2} - \dots - p_k E \quad (k = 1, 2, \dots, n-1). \quad (33)$$

The matrices  $B_1, B_2, \dots, B_{n-1}$  can be computed in succession, starting from the recurrence relation

$$B_k = A B_{k-1} - p_k E \quad (k = 1, 2, \dots, n-1; B_0 = E). \quad (34)$$

Moreover,<sup>4</sup>

$$A B_{n-1} - p_n E = O. \quad (35)$$

The relations (34) and (35) follow immediately from (20) if we equate the coefficients of equal powers of  $\lambda$  on both sides.

If  $A$  is non-singular, then

$$p_n = (-1)^{n-1} |A| \neq 0,$$

and it follows from (35) that

$$A^{-1} = \frac{1}{p_n} B_{n-1}. \quad (36)$$

Let  $\lambda_0$  be a characteristic value of  $A$ , so that  $\Delta(\lambda_0) = 0$ . Substituting the value  $\lambda_0$  in (20), we find:

$$(\lambda_0 E - A) B(\lambda_0) = O. \quad (37)$$

Let us assume that  $B(\lambda_0) \neq O$  and denote by  $b$  an arbitrary non-zero column of this matrix. Then from (37) we have  $(\lambda_0 E - A)b = O$  or

$$Ab = \lambda_0 b. \quad (38)$$

Therefore every non-zero column of  $B(\lambda_0)$  determines a characteristic vector corresponding to the characteristic value  $\lambda_0$ .<sup>5</sup>

Thus:

<sup>4</sup> From (34) follows (33). If we substitute in (35) the expression for  $B_{n-1}$  given in (33), we obtain  $\Delta(A) = 0$ . This approach to the Hamilton-Cayley Theorem does not require the Generalized Bézout Theorem explicitly, but contains this theorem implicitly.

<sup>5</sup> See Chapter III, § 7. If to the characteristic value  $\lambda_0$  there correspond  $d_0$  linearly independent characteristic vectors ( $n - d_0$  is the rank of  $\lambda_0 E - A$ ), then the rank of  $B(\lambda_0)$  does not exceed  $d_0$ . In particular, if only one characteristic direction corresponds to  $\lambda_0$ , then in  $B(\lambda_0)$  the elements of any two columns are proportional.



If the coefficients of the characteristic polynomial are known, then the adjoint matrix can be found by formula (31). If the given matrix  $A$  is non-singular, then the inverse matrix  $A^{-1}$  can be found by formula (36). If  $\lambda_0$  is a characteristic value of  $A$ , then the non-zero columns of  $B(\lambda_0)$  are characteristic vectors of  $A$  for  $\lambda = \lambda_0$ .

*Example.*

$$A = \begin{vmatrix} 2 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{vmatrix},$$

$$\Delta(\lambda) = |\lambda E - A| = \begin{vmatrix} \lambda - 2 & 1 & -1 \\ 0 & \lambda - 1 & -1 \\ 1 & -1 & \lambda - 1 \end{vmatrix} = \lambda^3 - 4\lambda^2 + 5\lambda - 2,$$

$$\delta(\lambda, \mu) = \frac{\Delta(\lambda) - \Delta(\mu)}{\lambda - \mu} = \lambda^2 + \lambda(\mu - 4) + \mu^2 - 4\mu + 5,$$

$$B(\lambda) = \delta(\lambda E, A) = \lambda^2 E + \lambda \underbrace{(A - 4E)}_{B_1} + \underbrace{A^2 - 4A + 5E}_{B_2}.$$

But

$$B_1 = A - 4E = \begin{vmatrix} -2 & -1 & 1 \\ 0 & -3 & 1 \\ -1 & 1 & -3 \end{vmatrix}, \quad B_2 = AB_1 + 5E = \begin{vmatrix} 0 & 2 & -2 \\ -1 & 3 & -2 \\ 1 & -1 & 2 \end{vmatrix},$$

$$B(\lambda) = \begin{vmatrix} \lambda^2 - 2\lambda & -\lambda + 2 & \lambda - 2 \\ -\lambda & \lambda^2 - 3\lambda + 3 & \lambda - 2 \\ -\lambda + 1 & \lambda - 1 & \lambda^2 - 3\lambda + 2 \end{vmatrix},$$

$$|A| = 2, \quad A^{-1} = \frac{1}{2} B_2 = \begin{vmatrix} 0 & 1 & -1 \\ -\frac{1}{2} & \frac{3}{2} & -1 \\ \frac{1}{2} & -\frac{1}{2} & 1 \end{vmatrix}.$$

Furthermore,

$$\Delta(\lambda) = (\lambda - 1)^2 (\lambda - 2).$$

The first column of the matrix  $B(+1)$  gives the characteristic vector  $(+1, +1, 0)$  for the characteristic value  $\lambda = 1$ .

The first column of the matrix  $B(+2)$  gives the characteristic vector  $(0, +1, +1)$  corresponding to the characteristic value  $\lambda = 2$ .

### § 5. The Method of Faddeev for the Simultaneous Computation of the Coefficients of the Characteristic Polynomial and of the Adjoint Matrix

1. D. K. Faddeev<sup>6</sup> has suggested a method for the simultaneous determination of the scalar coefficients  $p_1, p_2, \dots, p_n$  of the characteristic polynomial

$$\Delta(\lambda) = \lambda^n - p_1\lambda^{n-1} - p_2\lambda^{n-2} - \dots - p_n \quad (39)$$

and of the matrix coefficients  $B_1, B_2, \dots, B_{n-1}$  of the adjoint matrix  $B(\lambda)$ .

In order to explain the method of Faddeev<sup>7</sup> we introduce the concept of the trace (or spur) of a matrix.

By the *trace*  $\text{tr } A$  of a matrix  $A = \|a_{ik}\|_1^n$  we mean the sum of the diagonal elements of the matrix:

$$\text{tr } A = \sum_{i=1}^n a_{ii}. \quad (40)$$

It is easy to see that

$$\text{tr } A = p_1 = \sum_{i=1}^n \lambda_i, \quad (41)$$

if  $\lambda_1, \lambda_2, \dots, \lambda_n$  are the characteristic values of  $A$ , i.e., if

$$\Delta(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \dots (\lambda - \lambda_n). \quad (42)$$

Since by Theorem 3  $A^k$  has the characteristic values  $\lambda_1^k, \lambda_2^k, \dots, \lambda_n^k$  ( $k = 0, 1, 2, \dots$ ), we have

$$\text{tr } A^k = s_k = \sum_{i=1}^n \lambda_i^k \quad (k = 0, 1, 2, \dots). \quad (43)$$

The sums  $s_k$  ( $k = 1, 2, \dots, n$ ) of powers of the roots of the polynomial (39) are connected with the coefficients by Newton's formulas<sup>8</sup>

$$kp_k = s_k - p_1 s_{k-1} - \dots - p_{k-1} s_1 \quad (k = 1, 2, \dots, n). \quad (44)$$

If the traces  $s_1, s_2, \dots, s_n$  of the matrices  $A, A^2, \dots, A^n$  are computed, then the coefficients  $p_1, p_2, \dots, p_n$  can be determined from (44). This is the method of Leverrier for the determination of the coefficients of the characteristic polynomial from the traces of the powers of the matrix.

2. Faddeev has proposed to compute successively, instead of the traces of the powers  $A, A^2, \dots, A^n$ , the traces of certain other matrices  $A_1, A_2, \dots, A_n$

<sup>6</sup> See [14], p. 160.

<sup>7</sup> In Chapter VII, § 8, we shall discuss another effective method, due to A. N. Krylov, of computing the coefficients of the characteristic polynomial.

<sup>8</sup> See, for example, G. Chrystal, *Textbook of Algebra*, Vol. I, pp. 436ff.

and so to determine  $p_1, p_2, \dots, p_n$  and  $B_1, B_2, \dots, B_n$  by the following formulas:

$$\left. \begin{aligned} A_1 &= A, & p_1 &= \text{tr } A_1, & B_1 &= A_1 - p_1 E, \\ A_2 &= AB_1, & p_2 &= \frac{1}{2} \text{tr } A_2, & B_2 &= A_2 - p_2 E \\ &\dots\dots\dots & & & & \\ A_{n-1} &= AB_{n-2}, & p_{n-1} &= \frac{1}{n-1} \text{tr } A_{n-1}, & B_{n-1} &= A_{n-1} - p_{n-1} E, \\ A_n &= AB_{n-1}, & p_n &= \frac{1}{n} \text{tr } A_n, & B_n &= A_n - p_n E = O. \end{aligned} \right\} \quad (45)$$

The last equation  $B_n = A_n - p_n E = O$  may be used to check the computation.

In order to convince ourselves that the numbers  $p_1, p_2, \dots, p_n$  and the matrices  $B_1, B_2, \dots, B_{n-1}$  that are determined successively by (45) are, in fact, the coefficients of  $\Delta(\lambda)$  and  $B(\lambda)$ , we note that the following formulas for  $A_k$  and  $B_k$  ( $k = 1, 2, \dots, n$ ) follow from (45):

$$A_k = A^k - p_1 A^{k-1} - \dots - p_{k-1} A, \quad B_k = A^k - p_1 A^{k-1} - \dots - p_{k-1} A - p_k E. \quad (46)$$

Equating the traces on the left-hand and right-hand sides of the first of these formulas, we obtain

$$kp_k = s_k - p_1 s_{k-1} - \dots - p_{k-1} s_1.$$

But these formulas coincide with Newton's formulas (44) by which the coefficients of the characteristic polynomial  $\Delta(\lambda)$  are determined successively. Therefore the numbers  $p_1, p_2, \dots, p_n$  determined by (45) are also the coefficients of  $\Delta(\lambda)$ . But then the second of formulas (46) coincide with formulas (33) by which the matrix coefficients  $B_1, B_2, \dots, B_{n-1}$  of the adjoint matrix  $B(\lambda)$  are determined. Therefore, formulas (45) also determine the coefficients  $B_1, B_2, \dots, B_{n-1}$  of the matrix polynomial  $B(\lambda)$ .

*Example.*<sup>9</sup>

$$A = \begin{vmatrix} 2 & -1 & 1 & 2 \\ 0 & 1 & 1 & 0 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 2 & 2 & 4 & 3 \end{vmatrix}, \quad p_1 = \text{tr } A = 4, \quad B_1 = A - 4E = \begin{vmatrix} -2 & -1 & 1 & 2 \\ 0 & -3 & 1 & 0 \\ -1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -4 \end{vmatrix};$$

<sup>9</sup> As a check on the computation, we write under each matrix  $A_1, A_2, A_3$  a row whose elements are the sums of the elements above it. The product of this row of 'column-sums' of the first factor into the columns of the second factor must give the elements of the column-sum of the product.

# MINIMAL POLYNOMIAL OF MATRIX.

$$A_1 = AB_1 = \begin{vmatrix} -3 & 4 & 0 & -3 \\ -1 & -2 & -2 & 1 \\ 2 & 0 & -2 & -5 \\ -3 & -3 & -1 & 3 \\ -5 & -1 & -5 & -4 \end{vmatrix}, p_1 = \frac{1}{2} \operatorname{tr} A_1 = -2, B_1 = A_1 + 2E = \begin{vmatrix} -1 & 4 & 0 & -3 \\ -1 & 0 & -2 & 1 \\ 2 & 0 & 0 & -5 \\ -3 & -3 & -1 & 3 \end{vmatrix}.$$

$$A_2 = AB_2 = \begin{vmatrix} -5 & 2 & 0 & -2 \\ 1 & 0 & -2 & -4 \\ -1 & -7 & -3 & 4 \\ 0 & 4 & -2 & -7 \\ -5 & -1 & -7 & -9 \end{vmatrix}, p_2 = \frac{1}{3} \operatorname{tr} A_2 = -5, B_2 = A_2 + 5E = \begin{vmatrix} 0 & 7 & 5 \\ 1 & 5 & -2 \\ -1 & -7 & 2 \\ 0 & 4 & -2 & -2 \end{vmatrix}.$$

$$A_3 = AB_3 = \begin{vmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{vmatrix}, p_3 = -2.$$

$$\Delta(\lambda) = \lambda^4 - 4\lambda^3 + 2\lambda^2 + 5\lambda + 2,$$

$$|A| = 2, A^{-1} = \frac{1}{p_4} B_3 = \begin{vmatrix} 0 & -1 & 0 & 1 \\ -\frac{1}{2} & -\frac{5}{2} & 1 & -2 \\ \frac{1}{2} & \frac{7}{2} & -1 & -2 \\ 0 & -2 & 1 & 1 \end{vmatrix}.$$

*Note.* If we wish to determine  $p_1, p_2, p_3, p_4$  and only the first columns of  $B_1, B_2, B_3$ , it is sufficient to compute in  $A_1$  the elements of the first column and only the diagonal elements of the remaining columns, in  $A_2$  only the elements of the first column, and in  $A_3$  only the first two elements of the first column.

## § 6. The Minimal Polynomial of a Matrix

1. DEFINITION 1: A scalar polynomial  $f(\lambda)$  is called an annihilating polynomial of the square matrix  $A$  if

$$f(A) = O.$$

An annihilating polynomial  $\psi(\lambda)$  of least degree with highest coefficient 1 is called a *minimal polynomial* of  $A$ .

By the Hamilton-Cayley Theorem the characteristic polynomial  $\Delta(\lambda)$  is an annihilating polynomial of  $A$ . However, as we shall show below, it is not, in general, a minimal polynomial.

Let us divide an arbitrary annihilating polynomial  $f(\lambda)$  by a minimal polynomial

$$f(\lambda) = \psi(\lambda)q(\lambda) + r(\lambda),$$

where the degree of  $r(\lambda)$  is less than that of  $\psi(\lambda)$ . Hence we have:

$$f(A) = \psi(A)q(A) + r(A).$$

Since  $f(A) = 0$  and  $\psi(A) = 0$ , it follows that  $r(A) = 0$ . But the degree of  $r(\lambda)$  is less than that of the minimal polynomial  $\psi(\lambda)$ . Therefore  $r(\lambda) \equiv 0$ .<sup>10</sup> Hence: *Every annihilating polynomial of a matrix is divisible without remainder by the minimal polynomial.*

Let  $\psi_1(\lambda)$  and  $\psi_2(\lambda)$  be two minimal polynomials of one and the same matrix. Then each is divisible without remainder by the other, i.e., the polynomials differ by a constant factor. This constant factor must be 1, because the highest coefficients in  $\psi_1(\lambda)$  and  $\psi_2(\lambda)$  are 1. Thus we have proved the *uniqueness of the minimal polynomial* of a given matrix  $A$ .

2. We shall now derive a formula connecting the minimal polynomial with the characteristic polynomial.

We denote by  $D_{n-1}(\lambda)$  the greatest common divisor of all the minors of order  $n-1$  of the characteristic matrix  $\lambda E - A$ , i.e., of all the elements of the matrix  $B(\lambda) = \|b_{ik}(\lambda)\|_1^n$  (see the preceding section). Then

$$B(\lambda) = D_{n-1}(\lambda) C(\lambda), \quad (47)$$

where  $C(\lambda)$  is a certain polynomial matrix, the 'reduced' adjoint matrix of  $\lambda E - A$ . From (20) and (47) we have:

$$\Delta(\lambda) E = (\lambda E - A) C(\lambda) D_{n-1}(\lambda). \quad (48)$$

Hence it follows that  $\Delta(\lambda)$  is divisible without remainder by  $D_{n-1}(\lambda)$ .<sup>11</sup>

$$\frac{\Delta(\lambda)}{D_{n-1}(\lambda)} = \psi(\lambda), \quad (49)$$

where  $\psi(\lambda)$  is some polynomial. The factor  $D_{n-1}(\lambda)$  in (48) may be cancelled on both sides:<sup>12</sup>

$$\psi(\lambda) E = (\lambda E - A) C(\lambda). \quad (50)$$

<sup>10</sup> Otherwise there would exist an annihilating polynomial of degree less than that of the minimal polynomial.

<sup>11</sup> We could also verify this immediately by expanding the characteristic determinant  $\Delta(\lambda)$  with respect to the elements of an arbitrary row.

<sup>12</sup> In this case we have, apart from (50), also the identity (see (20'))

$$\psi(\lambda) E = C(\lambda) (\lambda E - A),$$

i.e.,  $C(\lambda)$  is at one and the same time the left quotient and right quotient of  $\psi(\lambda)E$  on division by  $\lambda E - A$ .

Since  $\psi(\lambda)E$  is divisible on the left without remainder by  $\lambda E - A$ , it follows by the Generalized Bézout Theorem that

$$\psi(A) = O.$$

Thus, the polynomial  $\psi(\lambda)$  defined by (49) is an annihilating polynomial of  $A$ . Let us show that it is the minimal polynomial.

We denote the minimal polynomial by  $\psi^*(\lambda)$ . Then  $\psi(\lambda)$  is divisible by  $\psi^*(\lambda)$  without remainder:

$$\psi(\lambda) = \psi^*(\lambda)\chi(\lambda). \quad (51)$$

Since  $\psi^*(A) = O$ , by the Generalized Bézout Theorem the matrix polynomial  $\psi^*(\lambda)E$  is divisible on the left by  $\lambda E - A$  without remainder:

$$\psi^*(\lambda)E = (\lambda E - A)C^*(\lambda). \quad (52)$$

From (51) and (52) it follows that

$$\psi(\lambda)E = (\lambda E - A)C^*(\lambda)\chi(\lambda). \quad (53)$$

The identities (50) and (53) show that  $C(\lambda)$  as well as  $C^*(\lambda)\chi(\lambda)$  are left quotients of  $\psi(\lambda)E$  on division by  $\lambda E - A$ . By the uniqueness of division

$$C(\lambda) = C^*(\lambda)\chi(\lambda).$$

Hence it follows that  $\chi(\lambda)$  is a common divisor of all the elements of the polynomial matrix  $C(\lambda)$ . But, on the other hand, the greatest common divisor of all the elements of the reduced adjoint matrix  $C(\lambda)$  is equal to 1, because the matrix was obtained from  $B(\lambda)$  by division by  $D_{n-1}(\lambda)$ . Therefore  $\chi(\lambda) = \text{const.}$  Since the highest coefficients of  $\psi(\lambda)$  and  $\psi^*(\lambda)$  are equal, we have in (51)  $\chi(\lambda) = 1$ , i.e.,  $\psi(\lambda) = \psi^*(\lambda)$ , and this is what we had to prove.

We have established the following formula for the minimal polynomial:

$$\psi(\lambda) = \frac{\Delta(\lambda)}{D_{n-1}(\lambda)}. \quad (54)$$

3. For the reduced adjoint matrix  $C(\lambda)$  we have a formula analogous to (31) (p. 84):

$$C(\lambda) = \Psi(\lambda E, A); \quad (55)$$

where the polynomial  $\Psi(\lambda, \mu)$  is defined by the equation<sup>13</sup>

<sup>13</sup> Formula (55) can be deduced in the same way as (31). On both sides of the identity  $\psi(\lambda) - \psi(\mu) = (\lambda - \mu)\Psi(\lambda, \mu)$  we substitute for  $\lambda$  and  $\mu$  the matrices  $\lambda E$  and  $A$  and compare the matrix equation so obtained with (50).

$$\Psi(\lambda, \mu) = \frac{\psi(\lambda) - \psi(\mu)}{\lambda - \mu}. \quad (56)$$

Moreover,

$$(\lambda E - A)C(\lambda) = \psi(\lambda)E. \quad (57)$$

Going over to determinants on both sides of (57), we obtain

$$\Delta(\lambda) |C(\lambda)| = [\psi(\lambda)]^n. \quad (58)$$

Thus,  $\Delta(\lambda)$  is divisible without remainder by  $\psi(\lambda)$  and some power of  $\psi(\lambda)$  is divisible without remainder by  $\Delta(\lambda)$ , i.e., the sets of all the distinct roots of the polynomials  $\Delta(\lambda)$  and  $\psi(\lambda)$  are equal. In other words: *All the distinct characteristic values of  $A$  are roots of  $\psi(\lambda)$ .*

If

$$\Delta(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \dots (\lambda - \lambda_s)^{n_s} \quad (59)$$

$$(\lambda_i \neq \lambda_j \text{ for } i \neq j; n_i > 0, i, j = 1, 2, \dots, s),$$

then

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_s)^{m_s}, \quad (60)$$

where

$$0 < m_k \leq n_k \quad (k = 1, 2, \dots, s). \quad (61)$$

4. We mention one further property of the matrix  $C(\lambda)$ . Let  $\lambda_0$  be an arbitrary characteristic value of  $A = \|a_{ik}\|_1^n$ . Then  $\psi(\lambda_0) = 0$  and therefore, by (57),

$$(\lambda_0 E - A)C(\lambda_0) = O. \quad (62)$$

Note that  $C(\lambda_0) \neq O$  always holds, for otherwise all the elements of the reduced adjoint matrix  $C(\lambda)$  would be divisible without remainder by  $\lambda - \lambda_0$ , and this is impossible.

We denote by  $c$  an arbitrary non-zero column of  $C(\lambda_0)$ . Then from (62)

$$(\lambda_0 E - A)c = o,$$

i.e.,

$$Ac = \lambda_0 c. \quad (63)$$

In other words, every non-zero column of  $C(\lambda_0)$  (and such a column always exists) determines a characteristic vector for  $\lambda = \lambda_0$ .

*Example.*

$$A = \begin{vmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{vmatrix},$$

$$\Delta(\lambda) = \begin{vmatrix} \lambda - 3 & 3 & -2 \\ 1 & \lambda - 5 & 2 \\ 1 & -3 & \lambda \end{vmatrix} = \lambda^3 - 8\lambda^2 + 20\lambda - 16 = (\lambda - 2)^2(\lambda - 4),$$

$$\delta(\lambda, \mu) = \frac{\Delta(\mu) - \Delta(\lambda)}{\mu - \lambda} = \mu^2 + \mu(\lambda - 8) + \lambda^2 - 8\lambda + 20,$$

$$\begin{aligned} B(\lambda) &= A^2 + (\lambda - 8)A + (\lambda^2 - 8\lambda + 20)E \\ &= \begin{vmatrix} 10 & -18 & 12 \\ -6 & 22 & -12 \\ -6 & 18 & -8 \end{vmatrix} + (\lambda - 8) \begin{vmatrix} 3 & -3 & 2 \\ -1 & 5 & -2 \\ -1 & 3 & 0 \end{vmatrix} + (\lambda^2 - 8\lambda + 20) \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \\ &= \begin{vmatrix} \lambda^2 - 5\lambda + 6 & -3\lambda + 6 & 2\lambda - 4 \\ -\lambda + 2 & \lambda^2 - 3\lambda + 2 & -2\lambda + 4 \\ -\lambda + 2 & 3\lambda - 6 & \lambda^2 - 8\lambda + 2 \end{vmatrix}. \end{aligned}$$

All the elements of the matrix  $B(\lambda)$  are divisible by  $D_2(\lambda) = \lambda - 2$ . Cancelling this factor, we have:

$$C(\lambda) = \begin{vmatrix} \lambda - 3 & -3 & 2 \\ -1 & \lambda - 1 & -2 \\ -1 & 3 & \lambda - 6 \end{vmatrix}$$

and

$$\psi(\lambda) = \frac{\Delta(\lambda)}{\lambda - 2} = (\lambda - 2)(\lambda - 4).$$

In  $C(\lambda)$  we substitute for  $\lambda$  the value  $\lambda_0 = 2$ :

$$C(2) = \begin{vmatrix} -1 & -3 & 2 \\ -1 & 1 & -2 \\ -1 & 3 & -4 \end{vmatrix}.$$

The first column gives us the characteristic vector  $(1, 1, 1)$  for  $\lambda_0 = 2$ . The second column gives us the characteristic vector  $(-3, 1, 3)$  for the same characteristic value  $\lambda_0 = 2$ . The third column is a linear combination of the first two.

Similarly, setting  $\lambda_0 = 4$ , we find from the first column of the matrix  $C(4)$  the characteristic vector  $(1, -1, -1)$  corresponding to the characteristic value  $\lambda_0 = 4$ .

The reader should note that  $\psi(\lambda)$  and  $C(\lambda)$  could have been determined by a different method.

To begin with, let us find  $D_2(\lambda)$ .  $D_2(\lambda)$  can only have 2 and 4 as its roots. For  $\lambda = 4$  the second order minor

$$\begin{vmatrix} 1 & \lambda - 5 \\ 1 & -3 \end{vmatrix} = -\lambda + 2$$

of  $\Delta(\lambda)$  does not vanish. Therefore  $D_2(4) \neq 0$ . For  $\lambda = 2$  the columns of  $\Delta(\lambda)$  become proportional. Therefore all the minors of order two in  $\Delta(\lambda)$



vanish for  $\lambda = 2 : D_2(\lambda) = 0$ . Since the minor to be computed is of the first degree,  $D_2(\lambda)$  cannot be divisible by  $(\lambda - 2)^2$ . Therefore

$$D_4(\lambda) = \lambda - 2.$$

Hence

$$\psi(\lambda) = \frac{\Delta(\lambda)}{\lambda - 2} = (\lambda - 2)(\lambda - 4) = \lambda^2 - 6\lambda + 8,$$

$$\psi(\lambda, \mu) = \frac{\psi(\mu) - \psi(\lambda)}{\mu - \lambda} = \mu + \lambda - 6,$$

$$C(\lambda) = \psi(\lambda E, A) = A + (\lambda - 6)E = \begin{vmatrix} \lambda - 3 & -3 & 2 \\ -1 & \lambda - 1 & -2 \\ -1 & 3 & \lambda - 6 \end{vmatrix}.$$

# CHAPTER V

## FUNCTIONS OF MATRICES

### § 1. Definition of a Function of a Matrix

1. Let  $A = \| a_{ik} \|_1^n$  be a square matrix and  $f(\lambda)$  a function of a scalar argument  $\lambda$ . We wish to define what is to be meant by  $f(A)$ , i.e., we wish to extend the function  $f(\lambda)$  to a matrix value of the argument.

We already know the solution of this problem in the simplest special case where  $f(\lambda) = \gamma_0 \lambda^l + \gamma_1 \lambda^{l-1} + \dots + \gamma_l$  is a polynomial in  $\lambda$ . In this case,  $f(A) = \gamma_0 A^l + \gamma_1 A^{l-1} + \dots + \gamma_l E$ . Starting from this special case, we shall obtain a definition of  $f(A)$  in the general case.

We denote by

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_s)^{m_s} \quad (1)$$

the minimal polynomial<sup>1</sup> of  $A$  (where  $\lambda_1, \lambda_2, \dots, \lambda_s$  are all the distinct characteristic values of  $A$ ). The degree of this polynomial is  $m = \sum_{k=1}^s m_k$ .

Let  $g(\lambda)$  and  $h(\lambda)$  be two polynomials such that

$$g(A) = h(A). \quad (2)$$

Then the difference  $d(\lambda) = g(\lambda) - h(\lambda)$ , as an annihilating polynomial for  $A$ , is divisible by  $\psi(\lambda)$  without remainder; we shall write this as follows:

$$g(\lambda) \equiv h(\lambda) \pmod{\psi(\lambda)}. \quad (3)$$

Hence by (1)

$$d(\lambda_k) = 0, \quad d'(\lambda_k) = 0, \quad \dots, \quad d^{(m_k-1)}(\lambda_k) = 0 \quad (k = 1, 2, \dots, s),$$

i.e.,

$$g(\lambda_k) = h(\lambda_k), \quad g'(\lambda_k) = h'(\lambda_k), \quad \dots, \quad g^{(m_k-1)}(\lambda_k) = h^{(m_k-1)}(\lambda_k) \quad (4)$$

( $k = 1, 2, \dots, s$ ).

---

<sup>1</sup> See Chapter IV, § 6.

The  $m$  numbers

$$f(\lambda_k), f'(\lambda_k), \dots, f^{(m_k-1)}(\lambda_k) \quad (k=1, 2, \dots, s) \quad (5)$$

will be called *the values of the function  $f(\lambda)$  on the spectrum of the matrix  $A$*  and the set of all these values will be denoted symbolically by  $f(A_A)$ . If for a function  $f(\lambda)$  the values (5) exist (i.e., have meaning), then we shall say that *the function  $f(\lambda)$  is defined on the spectrum of the matrix  $A$* .

Equation (4) shows that the polynomials  $g(\lambda)$  and  $h(\lambda)$  have the same values on the spectrum of  $A$ . In symbols:

$$g(A_A) = h(A_A).$$

Our argument is reversible: from (4) follows (3) and therefore (2).

Thus, given a matrix  $A$ , the values of the polynomial  $g(\lambda)$  on the spectrum of  $A$  determine the matrix  $g(A)$  completely, i.e., all polynomials  $g(\lambda)$  that assume the same values on the spectrum of  $A$  have one and the same matrix value  $g(A)$ .

We postulate that the definition of  $f(A)$  in the general case be subject to the same principle: *The values of the function  $f(\lambda)$  on the spectrum of the matrix  $A$  must determine  $f(A)$  completely, i.e., all functions  $f(\lambda)$  having the same values on the spectrum of  $A$  must have the same matrix value  $f(A)$ .*

But then it is obvious that for the general definition of  $f(A)$  it is sufficient to look for a polynomial<sup>2</sup>  $g(\lambda)$  that assumes the same values on the spectrum of  $A$  as  $f(\lambda)$  does and to set:

$$f(A) = g(A).$$

We are thus led to the following definition:

DEFINITION 1: *If the function  $f(\lambda)$  is defined on the spectrum of the matrix  $A$ , then*

$$f(A_A) = g(A_A),$$

*where  $g(\lambda)$  is an arbitrary polynomial that assumes on the spectrum of  $A$  the same values as does  $f(\lambda)$ :*

$$f(A) = g(A).$$

Among all the polynomials with complex coefficients that assume on the spectrum of  $A$  the same values as  $f(\lambda)$  there is one and only one polynomial

---

<sup>2</sup> It will be proved in § 2 that such an interpolation polynomial always exists and an algorithm for the computation of the coefficients of the interpolation polynomial of least degree will be given.

$r(\lambda)$  that is of degree less than  $m$ .<sup>3</sup> This polynomial  $r(\lambda)$  is uniquely determined by the interpolation conditions:

$$r(\lambda_k) = f(\lambda_k), \quad r'(\lambda_k) = f'(\lambda_k), \quad \dots, \quad r^{(m_k-1)}(\lambda_k) = f^{(m_k-1)}(\lambda_k) \quad (6) \\ (k = 1, 2, \dots, s).$$

The polynomial  $r(\lambda)$  is called the *Lagrange-Sylvester interpolation polynomial* for  $f(\lambda)$  on the spectrum of  $A$ . Definition 1 can also be formulated as follows:

**DEFINITION 1':** Let  $f(\lambda)$  be a function defined on the spectrum of a matrix  $A$  and  $r(\lambda)$  the corresponding Lagrange-Sylvester interpolation polynomial. Then

$$f(A) = r(A).$$

*Note.* If the minimal polynomial  $\psi(\lambda)$  of a matrix  $A$  has no multiple roots<sup>4</sup> (in (1)  $m_1 = m_2 = \dots = m_s = 1$ ;  $s = m$ ), then for  $f(A)$  to have a meaning it is sufficient that  $f(\lambda)$  be defined at the characteristic values  $\lambda_1, \lambda_2, \dots, \lambda_m$ . But if  $\psi(\lambda)$  has multiple roots, then for some characteristic values the derivatives of  $f(\lambda)$  up to a certain order (see (6)) must be defined as well.

**Example 1:** Let us consider the matrix<sup>5</sup>

$$H = \begin{pmatrix} \overbrace{0 & 1 & 0 & \dots & 0}^n \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix}.$$

Its minimal polynomial is  $\lambda^n$ . Therefore the values of  $f(\lambda)$  on the spectrum of  $H$  are the numbers  $f(0), f'(0), \dots, f^{(n-1)}(0)$ , and the polynomial  $r(\lambda)$  is of the form

$$r(\lambda) = f(0) + \frac{f'(0)}{1!} \lambda + \dots + \frac{f^{(n-1)}(0)}{(n-1)!} \lambda^{n-1}.$$

Therefore

<sup>3</sup> This polynomial is obtained from any other polynomial having the same spectral values by taking the remainder on division by  $\psi(\lambda)$  of that polynomial.

<sup>4</sup> In Chapter VI it will be shown that  $A$  is a matrix of simple structure (see Chapter III, § 8) in this case, and this case only.

<sup>5</sup> The properties of the matrix  $H$  were worked out in the example on pp. 13-14.

$$f(H) = f(0)E + \frac{f'(0)}{1!}H + \cdots + \frac{f^{(n-1)}(0)}{(n-1)!}H^{n-1} = \left\| \begin{array}{cccc} f(0) & \frac{f'(0)}{1!} & \cdots & \frac{f^{(n-1)}(0)}{(n-1)!} \\ 0 & f(0) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(0) \end{array} \right\|$$

*Example 2:* Let us consider the matrix

$$J = \left\| \begin{array}{cccc} \lambda_0 & \overbrace{1 \ 0 \ \cdots \ 0}^n \\ 0 & \lambda_0 & 1 & \cdots \ 0 \\ \vdots & \vdots & \vdots & \ddots \vdots \\ 0 & 0 & 0 & \cdots \ 1 \\ 0 & 0 & 0 & \cdots \ \lambda_0 \end{array} \right\|.$$

Note that  $J = \lambda_0 E + H$ , so that  $J - \lambda_0 E = H$ . The minimal polynomial of  $J$  is clearly  $(\lambda - \lambda_0)^n$ . The interpolation polynomial  $r(\lambda)$  of  $f(\lambda)$  is given by the equation

$$r(\lambda) = f(\lambda_0) + \frac{f'(\lambda_0)}{1!}(\lambda - \lambda_0) + \cdots + \frac{f^{(n-1)}(\lambda_0)}{(n-1)!}(\lambda - \lambda_0)^{n-1}$$

Therefore

$$\begin{aligned} f(J) &= r(J) = f(\lambda_0)E + \frac{f'(\lambda_0)}{1!}H + \cdots + \frac{f^{(n-1)}(\lambda_0)}{(n-1)!}H^{n-1} \\ &= \left\| \begin{array}{cccc} f(\lambda_0) & \frac{f'(\lambda_0)}{1!} & \cdots & \frac{f^{(n-1)}(\lambda_0)}{(n-1)!} \\ 0 & f(\lambda_0) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(\lambda_0) \end{array} \right\|. \end{aligned}$$

2. We mention two properties of functions of matrices.

1. If two matrices  $A$  and  $B$  are similar and  $T$  transforms  $A$  into  $B$ ,

$$B = T^{-1}AT,$$

then the matrices  $f(A)$  and  $f(B)$  are also similar and  $T$  transforms  $f(A)$  into  $f(B)$ ,

$$f(B) = T^{-1}f(A)T.$$

For two similar matrices have equal minimal polynomials,<sup>6</sup> so that  $f(\lambda)$  assumes the same values on the spectrum of  $A$  and of  $B$ . Therefore there exists an interpolation polynomial  $r(\lambda)$  such that  $f(A) = r(A)$  and  $f(B) = r(B)$ . But then it follows<sup>6</sup> from the equation  $r(B) = T^{-1}r(A)T$  that

$$f(B) = T^{-1}f(A)T.$$

2. If  $A$  is a quasi-diagonal matrix

$$A = \{A_1, A_2, \dots, A_u\},$$

then

$$f(A) = \{f(A_1), f(A_2), \dots, f(A_u)\}.$$

Let us denote by  $r(\lambda)$  the Lagrange-Sylvester interpolation polynomial of  $f(\lambda)$  on the spectrum of  $A$ . Then it is easy to see that

$$f(A) = r(A) = \{r(A_1), r(A_2), \dots, r(A_u)\}. \quad (7)$$

On the other hand, the minimal polynomial  $\psi(\lambda)$  of  $A$  is an annihilating polynomial for each of the matrices  $A_1, A_2, \dots, A_u$ . Therefore it follows from the equation

$$f(A_A) = r(A_A)$$

that

$$f(A_{A_1}) = r(A_{A_1}), \dots, f(A_{A_u}) = r(A_{A_u}).$$

Therefore

$$f(A_1) = r(A_1), \dots, f(A_u) = r(A_u),$$

and equation (7) can be written as follows:

$$f(A) = \{f(A_1), f(A_2), \dots, f(A_u)\}. \quad (8)$$

*Example 1:* If the matrix  $A$  is of simple structure

$$A = T \{\lambda_1, \lambda_2, \dots, \lambda_n\} T^{-1},$$

then

$$f(A) = T \{f(\lambda_1), f(\lambda_2), \dots, f(\lambda_n)\} T^{-1}.$$

$f(A)$  has meaning if the function  $f(\lambda)$  is defined at  $\lambda_1, \lambda_2, \dots, \lambda_n$ .

<sup>6</sup> From  $B = T^{-1}AT$  it follows that  $B^k = T^{-1}A^kT$  ( $k = 0, 1, 2, \dots$ ). Hence for every polynomial  $g(\lambda)$  we have  $g(B) = T^{-1}g(A)T$ . Therefore it follows from  $g(A) = O$  that  $g(B) = O$ , and vice versa.

*Example 2:* Let  $J$  be a matrix of the following quasi-diagonal form

$$J = \begin{pmatrix} \overbrace{\begin{matrix} \lambda_1 & 1 & 0 & \dots & 0 \\ 0 & \lambda_1 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_1 \end{matrix}}^{v_1} & & \\ & \ddots & \\ & & \overbrace{\begin{matrix} \lambda_u & 1 & 0 & \dots & 0 \\ 0 & \lambda_u & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & \lambda_u & 1 \\ 0 & 0 & 0 & \dots & 0 & \lambda_u \end{matrix}}^{v_u} \end{pmatrix}$$

All the elements in the non-diagonal blocks are zero. By (8) (see also the example on pp. 12-13),

$$f(J) = \begin{pmatrix} f(\lambda_1) \frac{f'(\lambda_1)}{1!} \dots \frac{f^{(v_1-1)}(\lambda_1)}{(v_1-1)!} & & \\ 0 & f(\lambda_1) & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \frac{f'(\lambda_1)}{1!} \\ 0 & 0 & \dots & \dots & f(\lambda_1) \end{pmatrix} \dots \begin{pmatrix} f(\lambda_u) \frac{f'(\lambda_u)}{1!} \dots \frac{f^{(v_u-1)}(\lambda_u)}{(v_u-1)!} & & \\ 0 & f(\lambda_u) & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \frac{f'(\lambda_u)}{1!} \\ 0 & 0 & \dots & \dots & f(\lambda_u) \end{pmatrix}$$

Here, as in the matrix  $J$ , all the elements in the non-diagonal blocks are also zero.<sup>7</sup>

<sup>7</sup> It will be established later (Chapter VI, § 6 or Chapter VII, § 7) that an arbitrary matrix  $A = \|a_{ik}\|_1^n$  is always similar to some matrix of the form  $J$ :  $A = TJT^{-1}$ . Therefore (see 1, on p. 98) we always have  $f(A) = Tf(J)T^{-1}$ .

## § 2. The Lagrange-Sylvester Interpolation Polynomial

1. To begin with, we consider the case in which the characteristic equation  $|\lambda E - A| = 0$  has no multiple roots. The roots of this equation—the characteristic values of the matrix  $A$ —will be denoted by  $\lambda_1, \lambda_2, \dots, \lambda_n$ . Then

$$\psi(\lambda) = |\lambda E - A| = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n),$$

and condition (6) can be written as follows:

$$r(\lambda_k) = f(\lambda_k) \quad (k = 1, 2, \dots, n).$$

In this case,  $r(\lambda)$  is the ordinary Lagrange interpolation polynomial for the function  $f(\lambda)$  at the points  $\lambda_1, \lambda_2, \dots, \lambda_n$ :

$$r(\lambda) = \sum_{k=1}^n \frac{(\lambda - \lambda_1) \cdots (\lambda - \lambda_{k-1})(\lambda - \lambda_{k+1}) \cdots (\lambda - \lambda_n)}{(\lambda_k - \lambda_1) \cdots (\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k+1}) \cdots (\lambda_k - \lambda_n)} f(\lambda_k).$$

By Definition 1'

$$f(A) = r(A) = \sum_{k=1}^n \frac{(A - \lambda_1 E) \cdots (A - \lambda_{k-1} E)(A - \lambda_{k+1} E) \cdots (A - \lambda_n E)}{(\lambda_k - \lambda_1) \cdots (\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k+1}) \cdots (\lambda_k - \lambda_n)} f(\lambda_k).$$

2. Let us assume now that the characteristic polynomial has multiple roots, but that the minimal polynomial, which is a divisor of the characteristic polynomial, has only simple roots:<sup>a</sup>

$$\psi(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_m).$$

In this case (as in the preceding one) all the exponents  $m_k$  in (1) are equal to 1, and the equation (6) takes the form

$$r(\lambda_k) = f(\lambda_k) \quad (k = 1, 2, \dots, m).$$

$r(\lambda)$  is again the ordinary Lagrange interpolation polynomial and

$$f(A) = \sum_{k=1}^m \frac{(A - \lambda_1 E) \cdots (A - \lambda_{k-1} E)(A - \lambda_{k+1} E) \cdots (A - \lambda_m E)}{(\lambda_k - \lambda_1) \cdots (\lambda_k - \lambda_{k-1})(\lambda_k - \lambda_{k+1}) \cdots (\lambda_k - \lambda_m)} f(\lambda_k).$$

3. We now consider the general case:

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_s)^{m_s} \quad (m_1 + m_2 + \cdots + m_s = m).$$

We represent the rational function  $\frac{r(\lambda)}{\psi(\lambda)}$ , where the degree of  $r(\lambda)$  is less than the degree of  $\psi(\lambda)$ , as a sum of partial fractions:

<sup>a</sup> See footnote 4.



$$\frac{r(\lambda)}{\psi(\lambda)} = \sum_{k=1}^s \left[ \frac{\alpha_{k1}}{(\lambda - \lambda_k)^{m_k}} + \frac{\alpha_{k2}}{(\lambda - \lambda_k)^{m_k-1}} + \dots + \frac{\alpha_{km_k}}{\lambda - \lambda_k} \right], \quad (9)$$

where  $\alpha_{kj}$  ( $j = 1, 2, \dots, m_k$ ;  $k = 1, 2, \dots, s$ ) are certain constants.

In order to determine the numerators  $\alpha_{kj}$  of the partial fractions we multiply both sides of (9) by  $(\lambda - \lambda_k)^{m_k}$  and denote by  $\psi_k(\lambda)$  the polynomial

$\frac{\psi(\lambda)}{(\lambda - \lambda_k)^{m_k}}$ . Then we obtain:

$$\begin{aligned} \frac{r(\lambda)}{\psi_k(\lambda)} &= \alpha_{k1} + \alpha_{k2}(\lambda - \lambda_k) + \dots + \alpha_{km_k}(\lambda - \lambda_k)^{m_k-1} + \\ &+ (\lambda - \lambda_k)^{m_k} \varrho_k(\lambda) \quad (k = 1, 2, \dots, s), \end{aligned} \quad (10)$$

where  $\varrho_k(\lambda)$  is a rational function, regular for  $\lambda = \lambda_k$ .\*

Hence

$$\left. \begin{aligned} \alpha_{k1} &= \left[ \frac{r(\lambda)}{\psi_k(\lambda)} \right]_{\lambda = \lambda_k}, \\ \alpha_{k2} &= \left[ \frac{r(\lambda)}{\psi_k(\lambda)} \right]'_{\lambda = \lambda_k} = r(\lambda_k) \left[ \frac{1}{\psi_k(\lambda)} \right]'_{\lambda = \lambda_k} + r'(\lambda_k) \frac{1}{\psi_k(\lambda_k)}, \dots \quad (k = 1, 2, \dots, s). \end{aligned} \right\} \quad (11)$$

Formulas (11) show that the numerators  $\alpha_{kj}$  on the right-hand side of (9) are expressible in terms of the values of the polynomial  $r(\lambda)$  on the spectrum of  $A$ , and these values are known: they are equal to the corresponding values of the function  $f(\lambda)$  and its derivatives. Therefore

$$\alpha_{k1} = \frac{f(\lambda_k)}{\psi_k(\lambda_k)}, \quad \alpha_{k2} = f(\lambda_k) \left[ \frac{1}{\psi_k(\lambda)} \right]'_{\lambda = \lambda_k} + f'(\lambda_k) \frac{1}{\psi_k(\lambda_k)}, \dots \quad (k = 1, 2, \dots, s). \quad (12)$$

Formulas (12) may be abbreviated as follows:

$$\alpha_{kj} = \frac{1}{(j-1)!} \left[ \frac{f(\lambda)}{\psi_k(\lambda)} \right]^{(j-1)}_{\lambda = \lambda_k} \quad (j = 1, 2, \dots, m_k; \quad k = 1, 2, \dots, s). \quad (13)$$

When all the  $\alpha_{kj}$  have been found, we can determine  $r(\lambda)$  from the following formula, which is obtained from (9) by multiplying both sides by  $\psi(\lambda)$ :

$$r(\lambda) = \sum_{k=1}^s [\alpha_{k1} + \alpha_{k2}(\lambda - \lambda_k) + \dots + \alpha_{km_k}(\lambda - \lambda_k)^{m_k-1}] \psi_k(\lambda). \quad (14)$$

In this formula the expression in brackets that multiplies  $\psi_k(\lambda)$  is by (13), equal to the sum of the first  $m_k$  terms of the Taylor expansion of  $f(\lambda)$  in powers of  $(\lambda - \lambda_k)$ .

\* I.e., that does not become infinite for  $\lambda = \lambda_k$ .

*Note.* The Lagrange-Sylvester interpolation polynomial can be obtained by a limiting process from the Lagrange interpolation polynomial.

Let

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_s)^{m_s} \quad (m = \sum_{k=1}^s m_k).$$

We denote the Lagrange interpolation polynomial constructed for the  $m$  points

$$\lambda_1^{(1)}, \lambda_1^{(2)}, \dots, \lambda_1^{(m_1)}; \lambda_2^{(1)}, \lambda_2^{(2)}, \dots, \lambda_2^{(m_2)}; \dots; \lambda_s^{(1)}, \lambda_s^{(2)}, \dots, \lambda_s^{(m_s)}$$

by

$$L(\lambda) = L\left(\begin{array}{c} \lambda_1^{(1)}, \dots, \lambda_1^{(m_1)}; \dots; \lambda_s^{(1)}, \dots, \lambda_s^{(m_s)}; \\ f(\lambda_1^{(1)}), \dots, f(\lambda_1^{(m_1)}); \dots; f(\lambda_s^{(1)}), \dots, f(\lambda_s^{(m_s)}) \end{array}; \lambda\right).$$

Then it is not difficult to show that the required Lagrange-Sylvester polynomial is determined by the formula

$$r(\lambda) = \lim_{\substack{\lambda_1^{(1)}, \dots, \lambda_1^{(m_1)} \rightarrow \lambda_1 \\ \dots \dots \dots \dots \dots \dots \dots \dots \\ \lambda_s^{(1)}, \dots, \lambda_s^{(m_s)} \rightarrow \lambda_s}} L(\lambda).$$

*Example:*

$$\psi(\lambda) = (\lambda - \lambda_1)^3 (\lambda - \lambda_2)^2 \quad (m = 5).$$

Then

$$\frac{r(\lambda)}{\psi(\lambda)} = \frac{\alpha}{(\lambda - \lambda_1)^3} + \frac{\beta}{\lambda - \lambda_1} + \frac{\gamma}{(\lambda - \lambda_2)^2} + \frac{\delta}{(\lambda - \lambda_2)^2} + \frac{\varepsilon}{\lambda - \lambda_2}.$$

Hence

$$r(\lambda) = [\alpha + \beta(\lambda - \lambda_1)](\lambda - \lambda_2)^3 + [\gamma + \delta(\lambda - \lambda_2) + \varepsilon(\lambda - \lambda_2)^2](\lambda - \lambda_1)^3$$

and therefore

$$r(A) = [\alpha E + \beta(A - \lambda_1 E)](A - \lambda_2 E)^3 + [\gamma E + \delta(A - \lambda_2 E) + \varepsilon(A - \lambda_2 E)^2](A - \lambda_1 E)^3.$$

$\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ , and  $\varepsilon$  can be found from the following formulas:

$$\begin{aligned} \alpha &= \frac{f(\lambda_1)}{(\lambda_1 - \lambda_2)^3}, & \beta &= -\frac{3}{(\lambda_1 - \lambda_2)^4} f(\lambda_1) + \frac{1}{(\lambda_1 - \lambda_2)^3} f'(\lambda_1), \\ \gamma &= \frac{f(\lambda_2)}{(\lambda_2 - \lambda_1)^2}, & \delta &= -\frac{2}{(\lambda_2 - \lambda_1)^3} f(\lambda_2) + \frac{1}{(\lambda_2 - \lambda_1)^2} f'(\lambda_2), \\ \varepsilon &= \frac{3}{(\lambda_2 - \lambda_1)^4} f(\lambda_2) - \frac{2}{(\lambda_2 - \lambda_1)^3} f'(\lambda_2) + \frac{1}{2} \frac{1}{(\lambda_2 - \lambda_1)^2} f''(\lambda_2). \end{aligned}$$

### § 3. Other Forms of the Definition of $f(A)$ . The Components of the Matrix $A$

1. Let us return to the formula (14) for  $r(\lambda)$ . When we substitute in (14) the expressions (12) for the coefficients  $\alpha$  and combine the terms that contain one and the same value of the function  $f(\lambda)$  or of one of its derivatives, we represent  $r(\lambda)$  in the form

$$r(\lambda) = \sum_{k=1}^s \left[ f(\lambda_k) \varphi_{k1}(\lambda) + f'(\lambda_k) \varphi_{k2}(\lambda) + \dots + f^{(m_k-1)}(\lambda_k) \varphi_{km_k}(\lambda) \right]. \quad (15)$$

Here  $\varphi_{kj}(\lambda)$  ( $j = 1, 2, \dots, m_k; k = 1, 2, \dots, s$ ) are easily computable polynomials in  $\lambda$  of degree less than  $m$ . These polynomials are completely determined when  $\psi(\lambda)$  is given and do not depend on the choice of the function  $f(\lambda)$ . The number of these polynomials is equal to the number of values of the function  $f(\lambda)$  on the spectrum of  $A$ , i.e., equal to  $m$  ( $m$  is the degree of the minimal polynomial  $\psi(\lambda)$ ). The functions  $\varphi_{kj}(\lambda)$  represent the Lagrange-Sylvester interpolation polynomial for the function whose values on the spectrum of  $A$  are all equal to zero with the exception of  $f^{(j-1)}(\lambda_k)$ , which is equal to 1.

All the polynomials  $\varphi_{kj}(\lambda)$  ( $j = 1, 2, \dots, m_k; k = 1, 2, \dots, s$ ) are linearly independent. For suppose that

$$\sum_{k=1}^s \sum_{j=1}^{m_k} c_{kj} \varphi_{kj}(\lambda) = 0.$$

Let us determine the interpolation polynomial  $r(\lambda)$  from the  $m$  conditions:

$$r^{(j-1)}(\lambda_k) = c_{kj} \quad (j = 1, 2, \dots, m_k; k = 1, 2, \dots, s). \quad (16)$$

Then by (15) and (16)

$$r(\lambda) = \sum_{k=1}^s \sum_{j=1}^{m_k} c_{kj} \varphi_{kj}(\lambda) = 0$$

and, therefore, by (16)

$$c_{kj} = 0 \quad (j = 1, 2, \dots, m_k; k = 1, 2, \dots, s).$$

From (15) we deduce the *fundamental formula for  $f(A)$* :

$$f(A) = \sum_{k=1}^s \left[ f(\lambda_k) Z_{k1} + f'(\lambda_k) Z_{k2} + \dots + f^{(m_k-1)}(\lambda_k) Z_{km_k} \right], \quad (17)$$

where

$$Z_{kj} = \varphi_{kj}(A) \quad (j = 1, 2, \dots, m_k; k = 1, 2, \dots, s). \quad (18)$$

The matrices  $Z_{kj}$  are completely determined when  $A$  is given and do not depend on the choice of the function  $f(\lambda)$ . On the right-hand side of (17) the function  $f(\lambda)$  is represented only by its values on the spectrum of  $A$ .

The matrices  $Z_{kj}$  ( $j = 1, 2, \dots, m_k; k = 1, 2, \dots, s$ ) will be called the *constituent matrices* or *components* of the given matrix  $A$ .

The components  $Z_{kj}$  are linearly independent.

For suppose that

$$\sum_{k=1}^s \sum_{j=1}^{m_k} c_{kj} Z_{kj} = O.$$

Then by (18)

$$\chi(A) = O, \quad (19)$$

where

$$\chi(\lambda) = \sum_{k=1}^s \sum_{j=1}^{m_k} c_{kj} \varphi_{kj}(\lambda). \quad (20)$$

Since by (20) the degree of  $\chi(\lambda)$  is less than  $m$ , the degree of the minimal polynomial  $\psi(\lambda)$ , it follows from (19) that

$$\chi(\lambda) = 0.$$

But then, since the  $m$  functions  $\varphi_{kj}(\lambda)$  are linearly independent, (20) implies that

$$c_{kj} = 0 \quad (j = 1, 2, \dots, m_k; k = 1, 2, \dots, s),$$

and this is what we had to prove.

2. From the linear independence of the constituent matrices  $Z_{kj}$  it follows, among other things, that none of these matrices can be zero. Let us also note that any two components  $Z_{kj}$  are permutable among each other and with  $A$ , because they are all scalar polynomials in  $A$ .

The formula (17) for  $f(A)$  is particularly convenient to use when it is necessary to deal with several functions of one and the same matrix  $A$ , or when the function  $f(\lambda)$  depends not only on  $\lambda$ , but also on some parameter  $t$ . In the latter case, the components  $Z_{kj}$  on the right-hand side of (17) do not depend on  $t$ , and the parameter  $t$  enters only into the scalar coefficients of the matrices.

In the example at the end of § 2, where  $\psi(\lambda) = (\lambda - \lambda_1)^2(\lambda - \lambda_2)^3$ , we may represent  $r(\lambda)$  in the form

$$r(\lambda) = f(\lambda_1) \varphi_{11}(\lambda) + f'(\lambda_1) \varphi_{12}(\lambda) + f(\lambda_2) \varphi_{21}(\lambda) + f'(\lambda_2) \varphi_{22}(\lambda) + f''(\lambda_2) \varphi_{23}(\lambda),$$

where

$$\begin{aligned}\varphi_{11}(\lambda) &= \left( \frac{\lambda - \lambda_2}{\lambda_1 - \lambda_2} \right)^3 \left[ 1 - \frac{3(\lambda - \lambda_1)}{\lambda_1 - \lambda_2} \right], & \varphi_{12}(\lambda) &= \frac{(\lambda - \lambda_1)(\lambda - \lambda_2)^2}{(\lambda_1 - \lambda_2)^3}, \\ \varphi_{21}(\lambda) &= \left( \frac{\lambda - \lambda_1}{\lambda_2 - \lambda_1} \right)^2 \left[ 1 - \frac{2(\lambda - \lambda_2)}{\lambda_2 - \lambda_1} + \frac{3(\lambda - \lambda_2)^2}{(\lambda_2 - \lambda_1)^2} \right], \\ \varphi_{22}(\lambda) &= \frac{(\lambda - \lambda_1)^2(\lambda - \lambda_2)}{(\lambda_2 - \lambda_1)^2} \left[ 1 - \frac{2(\lambda - \lambda_2)}{\lambda_2 - \lambda_1} \right], \\ \varphi_{23}(\lambda) &= \frac{(\lambda - \lambda_1)^2(\lambda - \lambda_2)^2}{2(\lambda_2 - \lambda_1)^3}.\end{aligned}$$

Therefore

$$f(A) = f(\lambda_1)Z_{11} + f'(\lambda_1)Z_{12} + f(\lambda_2)Z_{21} + f'(\lambda_2)Z_{22} + f''(\lambda_2)Z_{23}$$

where

$$\begin{aligned}Z_{11} &= \varphi_{11}(A) = \frac{1}{(\lambda_1 - \lambda_2)^3} (A - \lambda_2 E)^3 \left[ E - \frac{3}{\lambda_1 - \lambda_2} (A - \lambda_1 E) \right], \\ Z_{12} &= \varphi_{12}(A) = \frac{1}{(\lambda_1 - \lambda_2)^3} (A - \lambda_1 E) (A - \lambda_2 E)^2, \dots\end{aligned}$$

3. When the matrix  $A$  is given and its components have actually to be found, we can set in the fundamental formula (17)  $f(\mu) = \frac{1}{\lambda - \mu}$ , where  $\lambda$  is a parameter. Then we obtain

$$(\lambda E - A)^{-1} = \frac{C(\lambda)}{\psi(\lambda)} = \sum_{k=1}^s \left[ \frac{Z_{k1}}{\lambda - \lambda_k} + \frac{1! Z_{k2}}{(\lambda - \lambda_k)^2} + \dots + \frac{(m_k - 1)! Z_{km_k}}{(\lambda - \lambda_k)^{m_k}} \right], \quad (21)$$

where  $C(\lambda)$  is the reduced adjoint matrix of  $\lambda E - A$  (Chapter IV, § 6).<sup>10</sup>

The matrices  $(j-1)! Z_{kj}$  are the numerators of the partial fractions in the decomposition (21), and by analogy with (9) they may be expressed by the values of  $C(\lambda)$  on the spectrum of  $A$  by formulas similar to (11):

$$(m_k - 1)! Z_{km_k} = \frac{C(\lambda_k)}{\psi_k(\lambda)}, \quad (m_k - 2)! Z_{k, m_k - 1} = \left[ \frac{C(\lambda)}{\psi_k(\lambda)} \right]'_{\lambda = \lambda_k},$$

Hence

$$Z_{kj} = \frac{1}{(j-1)!(m_k - j)!} \left[ \frac{C(\lambda)}{\psi_k(\lambda)} \right]^{(m_k - j)}_{\lambda = \lambda_k} \quad (j = 1, 2, \dots, m_k; k = 1, 2, \dots, s). \quad (22)$$

When we replace the constituent matrices in (17) by their expressions (22), we can represent the fundamental formula (17) in the form

<sup>10</sup> For  $f(\mu) = \frac{1}{\lambda - \mu}$  we have  $f(A) = (\lambda E - A)^{-1}$ . For  $f(A) = r(A)$ , where  $r(\mu)$  is the Lagrange-Sylvester interpolation polynomial. From the fact that  $f(\mu)$  and  $r(\mu)$  coincide on the spectrum of  $A$  it follows that  $(\lambda - \mu)r(\mu)$  and  $(\lambda - \mu)f(\mu) = 1$  coincide on this spectrum. Hence  $(\lambda E - A)r(A) = (\lambda E - A)f(A) = E$ .

$$f(A) = \sum_{k=1}^s \frac{1}{(m_k-1)!} \left[ \frac{C(\lambda)}{\psi_k(\lambda)} f(\lambda) \right]_{\lambda=\lambda_k}^{(m_k-1)} \quad (23)$$

*Example 1:*<sup>11</sup>

$$A = \begin{vmatrix} 2 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{vmatrix} \begin{vmatrix} 2 \\ 2 \\ 1 \end{vmatrix}, \quad \lambda E - A = \begin{vmatrix} \lambda-2 & 1 & -1 \\ 0 & \lambda-1 & -1 \\ 1 & -1 & \lambda-1 \end{vmatrix}.$$

In this case  $\Delta(\lambda) = |\lambda E - A| = (\lambda-1)^2(\lambda-2)$ . Since the minor of the element in the first row and second column of  $\lambda E - A$  is equal to 1, we have  $D_2(\lambda) = 1$  and, therefore,

$$\psi(\lambda) = \Delta(\lambda) = (\lambda-1)^2(\lambda-2) = \lambda^3 - 4\lambda^2 + 5\lambda - 2,$$

$$\Psi(\lambda, \mu) = \frac{\psi(\mu) - \psi(\lambda)}{\mu - \lambda} = \mu^2 + (\lambda-4)\mu + \lambda^2 - 4\lambda + 5$$

and

$$\begin{aligned} C(\lambda) &= \Psi(\lambda E, A) = A^2 + (\lambda-4)A + (\lambda^2 - 4\lambda + 5)E \\ &= \begin{vmatrix} 3 & -2 & 2 \\ -1 & 2 & 2 \\ -3 & 3 & 1 \end{vmatrix} \begin{vmatrix} 3 \\ 3 \\ 1 \end{vmatrix} + (\lambda-4) \begin{vmatrix} 2 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{vmatrix} + (\lambda^2 - 4\lambda + 5) \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}. \end{aligned}$$

The fundamental formula has in this case the form

$$f(A) = f(1)Z_{11} + f'(1)Z_{12} + f(2)Z_{21}. \quad (24)$$

Setting  $f(\mu) = \frac{1}{\lambda-\mu}$ , we find:

$$(\lambda E - A)^{-1} = \frac{C(\lambda)}{\psi(\lambda)} = \frac{Z_{11}}{\lambda-1} + \frac{Z_{12}}{(\lambda-1)^2} + \frac{Z_{21}}{\lambda-2};$$

hence

$$Z_{11} = -C(1) \cdot C'(1), \quad Z_{12} = -C(1), \quad Z_{21} = C(2).$$

We now use the above expression for  $C(\lambda)$ , compute  $Z_{11}$ ,  $Z_{12}$ ,  $Z_{21}$ , and substitute the results obtained in (24).

<sup>11</sup> The elements of the sum column are printed in italics and are used for checking the computation. When we multiply the rows of  $A$  into the sum column of  $B$  we obtain the sum column of  $AB$ .

$$\begin{aligned}
 f(A) &= f(1) \begin{vmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{vmatrix} + f'(1) \begin{vmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{vmatrix} + f(2) \begin{vmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{vmatrix} \\
 &= \begin{vmatrix} f(1) + f'(1) & -f'(1) & f'(1) \\ f(1) + f'(1) - f(2) & -f'(1) + f(2) & f'(1) \\ f(1) - f(2) & -f(1) + f(2) & f(1) \end{vmatrix}. \quad (25)
 \end{aligned}$$

*Example 2:* Let us show that we can determine  $f(A)$  starting only from the fundamental formula. Again let

$$A = \begin{vmatrix} 2 & -1 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & 1 \end{vmatrix}, \quad \psi(\lambda) = (\lambda - 1)^2(\lambda - 2).$$

Then

$$f(A) = f(1)Z_1 + f'(1)Z_2 + f(2)Z_3. \quad (24')$$

In (24') we substitute for  $f(\lambda)$  in succession 1,  $\lambda - 1$ ,  $(\lambda - 1)^2$ :

$$\begin{aligned}
 Z_1 + Z_3 &= E = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}, \\
 Z_1 + Z_3 &= A - E = \begin{vmatrix} 1 & -1 & 1 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{vmatrix}, \\
 Z_3 &= (A - E)^2 = \begin{vmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{vmatrix}.
 \end{aligned}$$

Computing the third equation from the first two term by term, we can determine all the  $Z$ . Substituting in (24'), we obtain the expression for  $f(A)$ .

4. The examples we have analyzed illustrate three methods of practical computation of  $f(A)$ . In the first method, we found the interpolation polynomial  $r(\lambda)$  and put  $f(A) = r(A)$ . In the second method, we made use of the decomposition (21) and expressed the components  $Z_{kj}$  in (17) by the values of the reduced adjoint matrix  $C(\lambda)$  on the spectrum of  $A$ . In the third method, we started from the fundamental formula (17) and substituted in succession certain simple polynomials for  $f(\lambda)$ ; from the linear equations so obtained we determined the constituent matrices  $Z_{kj}$ .

The third method is perhaps the most convenient for practical purposes. In the general case it can be stated as follows:

In (17) we substitute for  $f(\lambda)$  successively certain polynomials  $g_1(\lambda)$ ,  $g_2(\lambda)$ , ...,  $g_m(\lambda)$ :

$$g_i(A) = \sum_{k=1}^s [g_i(\lambda_k) Z_{k1} + g'_i(\lambda_k) Z_{k2} + \dots + g_i^{(m_k-1)}(\lambda_k) Z_{k m_k}]$$

$$(i = 1, 2, \dots, m).$$
(26)

From the  $m$  equations (26) we determine the matrices  $Z_{kj}$  and substitute the expressions so obtained in (17).

The result of eliminating  $Z_{kj}$  from the  $(m+1)$  equations (26) and (17) can be written in the form

$$\begin{vmatrix} f(A) & f(\lambda_1) \dots f^{(m_1-1)}(\lambda_1) \dots f(\lambda_s) \dots f^{(m_s-1)}(\lambda_s) \\ g_1(A) & g_1(\lambda_1) \dots g_1^{(m_1-1)}(\lambda_1) \dots g_1(\lambda_s) \dots g_1^{(m_s-1)}(\lambda_s) \\ \vdots & \vdots \\ g_m(A) & g_m(\lambda_1) \dots g_m^{(m_1-1)}(\lambda_1) \dots g_m(\lambda_s) \dots g_m^{(m_s-1)}(\lambda_s) \end{vmatrix} = 0.$$

Expanding this determinant with respect to the elements of the first column, we obtain the required expression for  $f(A)$ . As the factor of  $f(A)$  we have here the determinant  $\Delta = |g_i^{(j)}(\lambda_k)|$  (in the  $i$ -th row of  $\Delta$  there are found the values of the polynomial  $g_i(\lambda)$  on the spectrum of  $A$ ;  $i = 1, 2, \dots, m$ ). In order to determine  $f(A)$  we must have  $\Delta \neq 0$ . This will be so if no linear combination<sup>12</sup> of the polynomials vanishes completely on the spectrum of  $A$ , i.e., is divisible by  $\psi(\lambda)$ .

The condition  $\Delta \neq 0$  is always satisfied when the degrees of the polynomial  $g_1(\lambda)$ ,  $g_2(\lambda)$ , ...,  $g_m(\lambda)$  are  $0, 1, \dots, m-1$ , respectively.<sup>13</sup>

5. In conclusion, we mention that high powers of a matrix  $A^n$  can be conveniently computed by formula (17) by setting  $f(\lambda)$  equal to  $\lambda^n$ .<sup>14</sup>

*Example:* Given the matrix  $A = \begin{vmatrix} 5 & -4 \\ 4 & -3 \end{vmatrix}$  it is required to compute the elements of  $A^{100}$ . The minimal polynomial of the matrix is  $\psi(\lambda) = (\lambda - 1)^2$ .

<sup>12</sup> With coefficients not all equal to zero.

<sup>13</sup> In the last example,  $m = 3$ ,  $g_1(\lambda) = 1$ ,  $g_2(\lambda) = \lambda - 1$ ,  $g_3(\lambda) = (\lambda - 1)^2$ .

<sup>14</sup> Formula (17) may also be used to compute the inverse matrix  $A^{-1}$ , by setting  $f(\lambda) = \frac{1}{\lambda}$  or, what is the same, by setting  $\lambda = 0$  in (21).



The fundamental formula is

$$f(A) = f(1)Z_1 + f'(1)Z_1.$$

Replacing  $f(\lambda)$  successively by 1 and  $\lambda - 1$ , we obtain:

$$Z_1 = E, \quad Z_2 = A - E.$$

Therefore

$$f(A) = f(1)E + f'(1)(A - E).$$

Setting  $f(\lambda) = \lambda^{100}$ , we find

$$A^{100} = E + 100(A - E) = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + 100 \begin{vmatrix} 4 & -4 \\ 4 & -4 \end{vmatrix} = \begin{vmatrix} 401 & -400 \\ 400 & -399 \end{vmatrix}.$$

#### § 4. Representation of Functions of Matrices by means of Series

1. Let  $A = \|a_{ik}\|_1^n$  be a matrix with the minimal polynomial (1):

$$\varphi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_s)^{m_s} \quad (m = \sum_{k=1}^s m_k).$$

Furthermore, let  $f(\lambda)$  be a function and let  $f_1(\lambda), f_2(\lambda), \dots, f_p(\lambda), \dots$  be a sequence of functions defined on the spectrum of  $A$ .

We shall say that the sequence of functions  $f_p(\lambda)$  converges for  $p \rightarrow \infty$  to some limit on the spectrum of  $A$  if the limits

$$\lim_{p \rightarrow \infty} f_p(\lambda_k), \quad \lim_{p \rightarrow \infty} f'_p(\lambda_k), \quad \dots, \quad \lim_{p \rightarrow \infty} f_p^{(mk-1)}(\lambda_k) \quad (k = 1, 2, \dots, s)$$

exist.

We shall say that the sequence of functions  $f_p(\lambda)$  converges for  $p \rightarrow \infty$  to the function  $f(\lambda)$  on the spectrum of  $A$ , and we shall write

$$\lim_{p \rightarrow \infty} f_p(A) = f(A)$$

if

$$\lim_{p \rightarrow \infty} f_p(\lambda_k) = f(\lambda_k), \quad \lim_{p \rightarrow \infty} f'_p(\lambda_k) = f'(\lambda_k), \quad \dots, \quad \lim_{p \rightarrow \infty} f_p^{(mk-1)}(\lambda_k) = f^{(mk-1)}(\lambda_k) \\ (k = 1, 2, \dots, s).$$

The fundamental formula

$$f(A) = \sum_{k=1}^s [f(\lambda_k)Z_{k1} + f'(\lambda_k)Z_{k2} + \dots + f^{(mk-1)}(\lambda_k)Z_{km_k}]$$

expresses  $f(A)$  in terms of the values of  $f(\lambda)$  on the spectrum of  $A$ . If we regard the matrix as a vector in a space  $\mathbf{R}_n$  of dimension  $n^2$ , then it follows from the fundamental formula, by the linear independence of the matrices  $Z_{kj}$ , that all the  $f(A)$  (for given  $A$ ) form an  $m$ -dimensional subspace of  $\mathbf{R}_n$ .

with basis  $Z_{kj}$  ( $j = 1, 2, \dots, m_k; k = 1, 2, \dots, s$ ). In this basis the 'vector'  $f(A)$  has as its coordinates the  $m$  values of the function  $f(\lambda)$  on the spectrum of  $A$ .

These considerations make the following theorem perfectly obvious:

**THEOREM 1:** *A sequence of matrices  $f_p(A)$  converges for  $p \rightarrow \infty$  to some limit if and only if the sequence  $f_p(\lambda)$  converges for  $p \rightarrow \infty$  on the spectrum of  $A$  to a limit, i.e., the limits*

$$\lim_{p \rightarrow \infty} f_p(A) \quad \text{and} \quad \lim_{p \rightarrow \infty} f_p(A_A)$$

*always exist simultaneously. Moreover, the equation*

$$\lim_{p \rightarrow \infty} f_p(A_A) = f(A_A) \quad (27)$$

*implies that*

$$\lim_{p \rightarrow \infty} f_p(A) = f(A) \quad (28)$$

*and conversely.*

*Proof.* 1) If the values of  $f_p(\lambda)$  converge on the spectrum of  $A$  for  $p \rightarrow \infty$  to limit values, then from the formulas

$$f_p(A) = \sum_{k=1}^s \left[ f_p(\lambda_k) Z_{k1} + f'_p(\lambda_k) Z_{k2} + \dots + f_p^{(m_k-1)}(\lambda_k) Z_{km_k} \right] \quad (29)$$

there follows the existence of the limit  $\lim_{p \rightarrow \infty} f_p(A)$ . On the basis of this formula and of (17) we deduce (28) from (27).

2) Suppose, conversely, that  $\lim_{p \rightarrow \infty} f_p(A)$  exists. Since the  $m$  constituent matrices  $Z$  are linearly independent, we can express, by (29), the  $m$  values of  $f_p(\lambda)$  on the spectrum of  $A$  (as a linear form) by the  $m$  elements of the matrix  $f_p(A)$ . Hence the existence of the limit  $\lim_{p \rightarrow \infty} f_p(A_A)$  follows, and (27) holds in the presence of (28).

According to this theorem, if a sequence of polynomials  $g_p(\lambda)$  ( $p = 1, 2, 3, \dots$ ) converges to the function  $f(\lambda)$  on the spectrum of  $A$ , then

$$\lim_{p \rightarrow \infty} g_p(A) = f(A).$$

**2.** This formula underlines the naturality and generality of our definition of  $f(A)$ .  $f(A)$  is always obtained from the  $g_p(A)$  by passing to the limit  $p \rightarrow \infty$ , provided only that the sequence of polynomials  $g_p(\lambda)$  converges to  $f(\lambda)$  on the spectrum of  $A$ . The latter condition is necessary for the existence of the limit  $\lim_{p \rightarrow \infty} g_p(A)$ .

We shall say that the series  $\sum_{p=0}^{\infty} u_p(\lambda)$  converges on the spectrum of  $A$  to the function  $f(\lambda)$  and we shall write

$$f(A) = \sum_{p=0}^{\infty} u_p(A), \quad (30)$$

if all the functions occurring here are defined on the spectrum of  $A$  and the following equations hold:

$$f(\lambda_k) = \sum_{p=0}^{\infty} u_p(\lambda_k), \quad f'(\lambda_k) = \sum_{p=0}^{\infty} u'_p(\lambda_k), \quad \dots, \quad f^{(m_k-1)}(\lambda_k) = \sum_{p=0}^{\infty} u_p^{(m_k-1)}(\lambda_k) \\ (k = 1, 2, \dots, s),$$

where the series on the right-hand sides of these equations converge. In other words, if we set

$$s_p(\lambda) = \sum_{q=0}^p u_q(\lambda) \quad (p = 0, 1, 2, \dots),$$

then (30) is equivalent to

$$f(A) = \lim_{p \rightarrow \infty} s_p(A). \quad (31)$$

It is obvious that the theorem just proved can be stated in the following equivalent form:

**THEOREM 1':** The series  $\sum_{p=0}^{\infty} u_p(A)$  converges to a matrix if and only if the series  $\sum_{p=0}^{\infty} u_p(\lambda)$  converges on the spectrum of  $A$ . Moreover, the equation

$$f(A) = \sum_{p=0}^{\infty} u_p(A)$$

implies that

$$f(A) = \sum_{p=0}^{\infty} u_p(A),$$

and conversely.

**3.** Suppose a power series is given with the circle of convergence  $|\lambda - \lambda_0| < R$  and the sum  $f(\lambda)$ :

$$f(\lambda) = \sum_{p=0}^{\infty} \alpha_p (\lambda - \lambda_0)^p \quad (|\lambda - \lambda_0| < R). \quad (32)$$

Since a power series may be differentiated term by term any number of times within the circle of convergence, (32) converges on the spectrum of any matrix whose characteristic values lie within the circle of convergence.

Thus we have:

**THEOREM 2:** *If the function  $f(\lambda)$  can be expanded in a power series in the circle  $|\lambda - \lambda_0| < r$ ,*

$$f(\lambda) = \sum_{p=0}^{\infty} \alpha_p (\lambda - \lambda_0)^p, \quad (33)$$

*then this expansion remains valid when the scalar argument  $\lambda$  is replaced by a matrix  $A$  whose characteristic values lie within the circle of convergence.*

*Note.* In this theorem we may allow a characteristic value  $\lambda_k$  of  $A$  to fall on the circumference of the circle of convergence; but we must then postulate in addition that the series (33), differentiated  $m_k - 1$  times term by term, should converge at the point  $\lambda = \lambda_k$ . It is well known that this already implies the convergence of the  $j$  times differentiated series (33) at the point  $\lambda_k$  to  $f^{(j)}(\lambda_k)$  for  $j = 0, 1, \dots, m_k - 1$ .

The theorem just proved leads, for example, to the following expansions:<sup>15</sup>

$$\begin{aligned} e^A &= \sum_{p=0}^{\infty} \frac{A^p}{p!}, \quad \cos A = \sum_{p=0}^{\infty} \frac{(-1)^p}{(2p)!} A^{2p}, \quad \sin A = \sum_{p=0}^{\infty} (-1)^p \frac{A^{2p+1}}{(2p+1)!}, \\ \cosh A &= \sum_{p=0}^{\infty} \frac{A^{2p}}{(2p)!}, \quad \sinh A = \sum_{p=0}^{\infty} \frac{A^{2p+1}}{(2p+1)!}, \\ (E - A)^{-1} &= \sum_{p=0}^{\infty} A^p \quad (|\lambda_k| < 1; k = 1, 2, \dots, s), \\ \ln A &= \sum_{p=1}^{\infty} \frac{(-1)^{p-1}}{p} (A - E)^p \quad (|\lambda_k - 1| < 1; k = 1, 2, \dots, s) \end{aligned}$$

(by  $\ln \lambda$  we mean here the so-called principal value of the many-valued function  $\text{Ln } \lambda$ , i.e., that branch for which  $\text{Ln } 1 = 0$ ).

Let  $G(u_1, u_2, \dots, u_i)$  be a polynomial in  $u_1, u_2, \dots, u_i$ ; let  $f_1(\lambda), f_2(\lambda), \dots, f_i(\lambda)$  be functions of  $\lambda$  defined on the spectrum of the matrix  $A$ , and let

$$g(\lambda) \equiv G[f_1(\lambda), f_2(\lambda), \dots, f_i(\lambda)].$$

Then from

$$g(A) = 0 \quad (34)$$

there follows:

$$G[f_1(A), f_2(A), \dots, f_i(A)] = 0. \quad (35)$$

<sup>15</sup> The expansions in the first two rows hold for an arbitrary matrix  $A$ .

For let us denote by  $f_1(\lambda), f_2(\lambda), \dots, f_l(\lambda)$  the Lagrange-Sylvester interpolation polynomials for  $r_1(\lambda), r_2(\lambda), \dots, r_l(\lambda)$ , and let us set:

$$G[f_1(A), f_2(A), \dots, f_l(A)] = G[r_1(A), r_2(A), \dots, r_l(A)] = h(A) = O,$$

Then (34) implies

$$h(\lambda) = G[r_1(\lambda), r_2(\lambda), \dots, r_l(\lambda)],$$

Hence it follows that

$$h(A) = 0. \quad (36)$$

and this is what we had to show.

This result allows us to extend identities between functions of a scalar variable to matrix values of the argument.

For example, from

$$\cos^2 \lambda + \sin^2 \lambda = 1$$

we obtain for an arbitrary matrix  $A$

$$\cos^2 A + \sin^2 A = E$$

(in this case  $G(u_1, u_2) = u_1^2 + u_2^2 - 1$ ,  $f_1(\lambda) = \cos \lambda$ , and  $f_2(\lambda) = \sin \lambda$ ).

Similarly, for every matrix  $A$

$$e^A e^{-A} = E,$$

i.e.,

$$e^{-A} = (e^A)^{-1}$$

Further, for every matrix  $A$

$$e^{iA} = \cos A + i \sin A$$

Let  $A$  be a non-singular matrix ( $|A| \neq 0$ ). We denote by  $\sqrt{\lambda}$  the single-valued branch of the many-valued function  $\sqrt{\lambda}$  that is defined in a domain not containing the origin and containing all the characteristic values of  $A$ . Then  $\sqrt{A}$  has a meaning. From  $(\sqrt{\lambda})^2 - \lambda = 0$  it now follows that

$$(\sqrt{A})^2 = A.$$

Let  $f(\lambda) = \frac{1}{\lambda}$  and let  $A = \|a_{ik}\|_1^n$  be a non-singular matrix. Then  $f(\lambda)$  is defined as the spectrum of  $A$ , and in the equation

$$\lambda f(\lambda) = 1$$

we can therefore replace  $\lambda$  by  $A$ :

$$A \cdot f(A) = E,$$

i.e.,<sup>18</sup>

$$f(A) = A^{-1}.$$

Denoting by  $r(\lambda)$  the interpolation polynomial for the function  $1/\lambda$  we may represent the inverse matrix  $A^{-1}$  in the form of a polynomial in  $A$ :

<sup>18</sup> We have already made use of this on p. 109. See footnote 10.

$$A^{-1} = r(A).$$

Let us consider a rational function  $\varrho(\lambda) = \frac{g(\lambda)}{h(\lambda)}$ , where  $g(\lambda)$  and  $h(\lambda)$  are co-prime polynomials in  $\lambda$ . This function is defined on the spectrum of  $A$  if and only if the characteristic values of  $A$  are not roots of  $h(\lambda)$ , i.e.,<sup>17</sup> if  $|h(A)| \neq 0$ . Under this assumption we may replace  $\lambda$  by  $A$  in the identity

$$\varrho(\lambda) h(\lambda) = g(\lambda),$$

obtaining:

$$\varrho(A) h(A) = g(A).$$

Hence

$$\varrho(A) = g(A) [h(A)]^{-1} = [h(A)]^{-1} g(A). \quad (37)$$

*Notes.* 1) If  $A$  is a linear operator in an  $n$ -dimensional space  $R$ , then  $f(A)$  is defined exactly like  $f(\lambda)$ :

$$f(A) = r(A),$$

where  $r(\lambda)$  is the Lagrange-Sylvester interpolation polynomial for  $f(\lambda)$  on the spectrum of the operator  $A$  (the spectrum of  $A$  is determined by the minimal annihilating polynomial  $\psi(\lambda)$  of  $A$ ).

According to this definition, if the matrix  $A = \|a_{ik}\|_1^n$  corresponds to the operator  $A$  in some basis of the space, then in the same basis the matrix  $f(A)$  corresponds to the operator  $f(A)$ . All the statements of this chapter in which there occurs a matrix  $A$  remain valid after replacement of the matrix  $A$  by the operator  $A$ .

2) We can also define<sup>18</sup> a function of a matrix  $f(A)$  starting from the characteristic polynomial

$$\Delta(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{n_k}$$

instead of the minimal polynomial

$$\psi(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{m_k}.$$

<sup>17</sup> See (25) on p. 84.

<sup>18</sup> See, for example, MacMillan, W. D., *Dynamics of Rigid Bodies* (New York, 1936).



$$\frac{dx}{dt} = Ax. \quad (39)$$

Here, and in what follows, we mean by the *derivative of a matrix* that matrix which is obtained from the given one by replacing all its elements by their derivatives. Therefore  $\frac{dx}{dt}$  is the column matrix with the elements  $\frac{dx_1}{dt}, \frac{dx_2}{dt}, \dots, \frac{dx_n}{dt}$ .

We shall seek a solution of the system of differential equations satisfying the following initial conditions:

$$x_1|_{t=0} = x_{10}, \quad x_2|_{t=0} = x_{20}, \quad \dots, \quad x_n|_{t=0} = x_{n0}$$

or, briefly,

$$x|_{t=0} = x_0. \quad (40)$$

Let us expand the unknown column  $x$  into a MacLaurin series in powers of  $t$ :

$$x = x_0 + \dot{x}_0 t + \ddot{x}_0 \frac{t^2}{2!} + \dots \left( \dot{x}_0 = \frac{dx}{dt} \Big|_{t=0}, \ddot{x}_0 = \frac{d^2x}{dt^2} \Big|_{t=0}, \dots \right). \quad (41)$$

Then by successive differentiations we find from (39):

$$\frac{d^2x}{dt^2} = A \frac{dx}{dt} = A^2x, \quad \frac{d^3x}{dt^3} = A \frac{d^2x}{dt^2} = A^3x, \quad \dots \quad (42)$$

Substituting the value  $t=0$  in (39) and (42), we obtain:

$$\dot{x}_0 = Ax_0, \quad \ddot{x}_0 = A^2x_0, \quad \dots$$

Now the series (41) can be written as follows:

$$x = x_0 + tAx_0 + \frac{t^2}{2!}A^2x_0 + \dots = e^{At}x_0. \quad (43)$$

By direct substitution in (39) we see<sup>21</sup> that (43) is a solution of the differential equation (39). Setting  $t=0$  in (43), we find:

$$x|_{t=0} = x_0.$$

Thus, the formula (43) gives us the solution of the given system of differential equations satisfying the initial conditions (40).

Let us set  $f(\lambda) = e^{\lambda t}$  in (17). Then

$$e^{At} = \|q_{ik}(t)\|_1^n = \sum_{k=1}^n (Z_{k1} + Z_{k2}t + \dots + Z_{km_k}t^{m_k-1}) e^{\lambda_k t}. \quad (44)$$

---

<sup>21</sup>  $\frac{d}{dt}(e^{At}) = \frac{d}{dt}\left(A + At + \frac{A^2t^2}{2!} + \dots\right) = A + A^2t + \frac{A^3t^2}{2!} + \dots = Ae^{At}.$



The solution (43) may then be written in the following form:

$$\left. \begin{aligned} x_1 &= q_{11}(t) x_{10} + q_{12}(t) x_{20} + \dots + q_{1n}(t) x_{n0} \\ x_2 &= q_{21}(t) x_{10} + q_{22}(t) x_{20} + \dots + q_{2n}(t) x_{n0} \\ &\vdots \\ x_n &= q_{n1}(t) x_{10} + q_{n2}(t) x_{20} + \dots + q_{nn}(t) x_{n0}, \end{aligned} \right\} \quad (45)$$

where  $x_{10}, x_{20}, \dots, x_{n0}$  are constants equal to the initial values of the unknown functions  $x_1, x_2, \dots, x_n$ .

Thus, the integration of the given system of differential equations reduces to the computation of the elements of the matrix  $e^{At}$ .

If  $t = t_0$  is taken as the initial value of the argument, then (43) is to be replaced by the formula

$$x = e^{A(t-t_0)} x_0. \quad (46)$$

*Example.*

$$\begin{aligned}\frac{dx_1}{dt} &= 3x_1 - x_2 + x_3, \\ \frac{dx_2}{dt} &= 2x_1 - x_3, \\ \frac{dx_3}{dt} &= x_1 - x_2 + 2x_3.\end{aligned}$$

The coefficient matrix is

$$A = \begin{vmatrix} 3 & -1 & 1 \\ 2 & 0 & 1 \\ 1 & -1 & 2 \end{vmatrix}.$$

We form the characteristic determinant

$$\Delta(\lambda) = - \begin{vmatrix} 3-\lambda & 1 & 1 \\ 2 & -\lambda & 1 \\ 1 & -1 & 2-\lambda \end{vmatrix} = (\lambda-1)(\lambda-2)^2.$$

The greatest common divisor of the minors of order 2 is  $D_2(\lambda) = 1$ . Therefore

$$\psi(\lambda) = \Delta(\lambda) = (\lambda - 1)(\lambda - 2)^2.$$

The fundamental formula is

$$f(A) = f(1)Z_1 + f(2)Z_2 + f'(2)Z_3.$$

For  $f(\lambda)$  we choose in succession  $1, \lambda - 2, (\lambda - 2)^2$ . We obtain:

$$\begin{aligned}
Z_1 + Z_2 &= E = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}, \\
-Z_1 + Z_3 &= A - 2E = \begin{vmatrix} 1 & -1 & 1 \\ 2 & -2 & 1 \\ 1 & -1 & 0 \end{vmatrix}, \\
Z_1 &= (A - 2E)^2 = \begin{vmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{vmatrix}.
\end{aligned}$$

Hence we determine  $Z_1$ ,  $Z_2$ ,  $Z_3$  and substitute in the fundamental formula

$$f(A) = f(1) \begin{vmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{vmatrix} + f(2) \begin{vmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{vmatrix} + f'(2) \begin{vmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{vmatrix}.$$

If we now replace  $f(\lambda)$  by  $e^{\lambda t}$ , we obtain:

$$e^{At} = e^t \begin{vmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 1 & 0 \end{vmatrix} + e^{2t} \begin{vmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & -1 & 1 \end{vmatrix} + te^{2t} \begin{vmatrix} 1 & -1 & 1 \\ 1 & -1 & 1 \\ 0 & 0 & 0 \end{vmatrix} = \begin{vmatrix} (1+t)e^{2t} & -te^{2t} & te^{2t} \\ -e^t + (1+t)e^{2t} & e^t - te^{2t} & te^{2t} \\ -e^t + e^{2t} & e^t - e^{2t} & e^{2t} \end{vmatrix}.$$

Thus

$$\begin{aligned}
x_1 &= C_1(1+t)e^{2t} - C_2te^{2t} + C_3te^{2t}, \\
x_2 &= C_1[-e^t + (1+t)e^{2t}] + C_2(e^t - te^{2t}) + C_3te^{2t}, \\
x_3 &= C_1(-e^t + e^{2t}) + C_2(e^t - e^{2t}) + C_3e^{2t},
\end{aligned}$$

where

$$C_1 = x_{10}, \quad C_2 = x_{20}, \quad C_3 = x_{30}.$$

2. We now consider a system of inhomogeneous linear differential equations with constant coefficients:

$$\left. \begin{aligned}
\frac{dx_1}{dt} &= a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + f_1(t) \\
\frac{dx_2}{dt} &= a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + f_2(t) \\
&\dots\dots\dots \\
\frac{dx_n}{dt} &= a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + f_n(t),
\end{aligned} \right\} \quad (47)$$

where  $f_i(t)$  ( $i=1, 2, \dots, n$ ) are continuous functions in the interval  $t_0 \leq t \leq t_1$ . Denoting by  $f(t)$  the column matrix with the elements  $f_1(t), f_2(t), \dots, f_n(t)$  and again setting  $A = \|a_{ik}\|_1^n$ , we write the system (47) as follows:

$$\frac{dx}{dt} = Ax + f(t). \quad (48)$$

We replace  $x$  by a new column  $z$  of unknown functions, connected with  $x$  by the relation

$$x = e^{At}z \quad (49)$$

Differentiating (49) term by term and substituting the expression for  $\frac{dx}{dt}$  in (48) we find:<sup>22</sup>

$$e^{At} \frac{dz}{dt} = f(t). \quad (50)$$

Hence<sup>23</sup>

$$z(t) = c + \int_{t_0}^t e^{-A\tau} f(\tau) d\tau \quad (51)$$

and so by (49)

$$x = e^{At} \left[ c + \int_{t_0}^t e^{-A\tau} f(\tau) d\tau \right] = e^{At} c + \int_{t_0}^t e^{A(t-\tau)} f(\tau) d\tau; \quad (52)$$

where  $c$  is a column with arbitrary constant elements.

When we give to the argument  $t$  in (52) the value  $t_0$  we find  $c = e^{-At_0} x_0$ ; so that (52) can be written in the following form:

$$x = e^{A(t-t_0)} x_0 + \int_{t_0}^t e^{A(t-\tau)} f(\tau) d\tau. \quad (53)$$

<sup>22</sup> See footnote 21.

<sup>23</sup> If a matrix function of a scalar argument is given,  $B(\tau) = \|b_{ik}(\tau)\|$  ( $i=1, 2, \dots, m$ ;  $k=1, 2, \dots, n$ ;  $t_1 \leq \tau \leq t_2$ ), then the integral  $\int_{t_1}^{t_2} B(\tau) d\tau$  is defined in the natural way:

$$\int_{t_1}^{t_2} B(\tau) d\tau = \left\| \int_{t_1}^{t_2} b_{ik}(\tau) d\tau \right\| \quad (i=1, 2, \dots, m; k=1, 2, \dots, n).$$



Substituting for  $e^{At}$  the series

$$E + A \frac{t}{1!} + A^2 \frac{t^2}{2!} + \dots$$

and replacing  $A$  by its expression from (56), we have:

$$r = r_0 + v_0 t + \frac{1}{2} g t^2 - \omega \times \left( v_0 t^2 + \frac{1}{3} g t^3 \right) + \omega \times \left[ \omega \times \left( \frac{2}{3} v_0 t^3 + \frac{1}{6} g t^4 \right) \right] + \dots$$

Considering that the angular velocity  $\omega$  is small (for the earth,  $\omega \approx 7.3 \times 10^{-5} \text{ sec}^{-1}$ ), we neglect the terms containing the second and higher powers of  $\omega$ ; for the additional displacement of the point due to the motion of the earth we then obtain the approximate formula

$$d = -\omega \times \left( v_0 t^2 + \frac{1}{3} g t^3 \right).$$

Returning to the exact solution (58), let us compute  $e^{At}$ . As a preliminary we establish that the minimal polynomial of the operator  $A$  has the form

$$\psi(\lambda) = \lambda(\lambda^2 + 4\omega^2).$$

For we find from (56)

$$\begin{aligned} A^2 x &= 4\omega \times (\omega \times x) = 4(\omega x) \omega - 4\omega^2 x, \\ A^3 x &= -2\omega \times A^2 x = 8\omega^3 (\omega \times x). \end{aligned}$$

Hence and from (56) it follows that the operators  $E, A, A^2$  are linearly independent and that

$$A^3 + 4\omega^2 A = O.$$

The minimal polynomial  $\psi(\lambda)$  has the simple roots  $0, 2\omega i, -2\omega i$ . The Lagrange interpolation formula for  $e^{At}$  has the form

$$1 + \frac{\sin 2\omega t}{2\omega} \lambda + \frac{1 - \cos 2\omega t}{4\omega^2} \lambda^2.$$

Then

$$e^{At} = E + \frac{\sin 2\omega t}{2\omega} A + \frac{1 - \cos 2\omega t}{4\omega^2} A^2.$$

Substituting this expression for  $e^{At}$  in (58) and replacing the operator  $A$  by its expression from (56), we find

$$\begin{aligned} r = r_0 + v_0 t + g \frac{t^2}{2} - \omega \times \left( \frac{1 - \cos 2\omega t}{2\omega^2} v_0 + \frac{2\omega t - \sin 2\omega t}{4\omega^3} g \right) + \\ + \omega \times \left[ \omega \times \left( \frac{2\omega t - \sin 2\omega t}{2\omega^3} v_0 + \frac{-1 + 2\omega^2 t^2 + \cos 2\omega t}{4\omega^4} g \right) \right]. \end{aligned} \quad (59)$$



$$\left. \begin{aligned} \cos(\sqrt{A}t) &= E - \frac{1}{2!}At^2 + \frac{1}{4!}A^2t^4 - \dots, \\ (\sqrt{A})^{-1}\sin(\sqrt{A}t) &= Et - \frac{1}{3!}At^3 + \frac{1}{5!}A^2t^5 - \dots \end{aligned} \right\} \quad (62)$$

Formula (61) comprises all solutions of the system (60) or (60'), as the initial values  $x_0$  and  $\dot{x}_0$  may be chosen arbitrarily.

The right-hand sides of the formulas (62) have a meaning even when  $|A| = 0$ . Therefore (61) is the general solution of the given system of differential equations also when  $|A| = 0$ , provided only that the functions  $\cos(\sqrt{A}t)$  and  $(\sqrt{A})^{-1}\sin(\sqrt{A}t)$ , which are part of this expression, are interpreted as the right-hand sides of the formulas (62).

We leave it to the reader to verify that the general solution of the inhomogeneous system

$$\frac{d^2x}{dt^2} + Ax = f(t) \quad (63)$$

satisfying the initial conditions  $x|_{t=0} = x_0$  and  $\frac{dx}{dt}|_{t=0} = \dot{x}_0$  can be written in the form

$$\begin{aligned} x &= \cos(\sqrt{A}t)x_0 + (\sqrt{A})^{-1}\sin(\sqrt{A}t)\dot{x}_0 + \\ &+ (\sqrt{A})^{-1}\int_0^t \sin[\sqrt{A}(t-\tau)]f(\tau)d\tau. \end{aligned} \quad (64)$$

If  $t = t_0$  is taken as the initial time, then in (61) and (64)  $\cos(\sqrt{A}t)$  and  $\sin(\sqrt{A}t)$  must be replaced by  $\cos(\sqrt{A}(t-t_0))$  and  $\sin(\sqrt{A}(t-t_0))$ , and  $\int_0^t$  by  $\int_{t_0}^t$ .

In the special case

$$f(t) = h \sin(pt + \alpha)$$

( $h$  is a constant column, and  $p$  and  $\alpha$  are numbers), (64) can be replaced by:

$$x = \cos(\sqrt{A}t)c + (\sqrt{A})^{-1}\sin(\sqrt{A}t)d + (A - p^2E)^{-1}h \sin(pt + \alpha),$$

where  $c$  and  $d$  are columns with arbitrary constant elements. This formula has meaning when  $p^2$  is not a characteristic value of the matrix  $A$  ( $|A - p^2E| \neq 0$ ).

### § 6. Stability of Motion in the Case of a Linear System

1. Let  $x_1, x_2, \dots, x_n$  be parameters that characterize the displacement of 'perturbed' motion of a given mechanical system from an original motion,<sup>27</sup> and suppose that these parameters satisfy a system of differential equations of the first order:

$$\frac{dx_i}{dt} = f_i(x_1, x_2, \dots, x_n, t) \quad (i=1, 2, \dots, n); \quad (65)$$

the independent variable  $t$  in these equations is the time, and the right-hand sides  $f_i(x_1, x_2, \dots, x_n, t)$  are continuous functions of the variables  $x_1, \dots, x_n$  in some domain containing the point  $x_1=0, x_2=0, \dots, x_n=0$  for all  $t \geq t_0$  ( $t_0$  is the initial time).

We now introduce the definition of stability of motion according to Lyapunov.<sup>28</sup>

The motion to be investigated is called *stable* if for every  $\varepsilon > 0$  we can find a  $\delta > 0$  such that for arbitrary initial values of the parameters  $x_{10}, x_{20}, \dots, x_{n0}$  (for  $t=t_0$ ) with moduli less than  $\delta$  the parameters  $x_1, x_2, \dots, x_n$  remain of moduli less than  $\varepsilon$  for the whole time of the motion ( $t \geq t_0$ ), i.e., if for every  $\varepsilon > 0$  we can find a  $\delta > 0$  such that from

$$|x_{i0}| < \delta \quad (i=1, 2, \dots, n) \quad (66)$$

it follows that

$$|x_i(t)| < \varepsilon \quad (t \geq t_0). \quad (67)$$

If, in addition, for some  $\delta > 0$  we always have  $\lim_{t \rightarrow +\infty} x_i(t) = 0$  ( $i=1, 2, \dots, n$ ) as long as  $|x_{i0}| < \delta$  ( $i=1, 2, \dots, n$ ), then the motion is called *asymptotically stable*.

We now consider a linear system, i.e., that special case when (65) is a system of linear homogeneous differential equations

$$\frac{dx_i}{dt} = \sum_{k=1}^n p_{ik}(t) x_k, \quad (68)$$

where the  $p_{ik}(t)$  are continuous functions for  $t \geq t_0$  ( $i, k=1, 2, \dots, n$ ).

In matrix form the system (68) can be written as follows:

<sup>27</sup> In these parameters, the motion to be studied is characterized by constant zero values  $x_1=0, x_2=0, \dots, x_n=0$ . Therefore in the mathematical treatment of the problem we speak of the 'stability' of the zero solution of the system (65) of differential equations.

<sup>28</sup> See [14], p. 13; [9], pp. 10-11; or [36], pp. 11-12. See also [3].



$$\frac{dx}{dt} = P(t)x, \quad (68')$$

where  $x$  is the column matrix with the elements  $x_1, x_2, \dots, x_n$  and  $P(t) = \|p_{ij}(t)\|_1^n$  is the coefficient matrix.

We denote by

$$q_{1j}(t), q_{2j}(t), \dots, q_{nj}(t) \quad (j=1, 2, \dots, n) \quad (69)$$

$n$  linearly independent solutions of (68).<sup>29</sup> The matrix  $Q(t) = \|q_{ij}\|_1^n$  whose columns are these solutions is called an *integral matrix* of the system (68).

Every solution of the system of linear homogeneous differential equations is obtained as a linear combination of  $n$  linearly independent solutions with constant coefficients:

$$x_i = \sum_{j=1}^n c_j q_{ij}(t) \quad (i=1, 2, \dots, n),$$

or in matrix form,

$$x = Q(t)c, \quad (70)$$

where  $c$  is the column matrix whose elements are arbitrary constants  $c_1, c_2, \dots, c_n$ .

We now choose the special integral matrix for which

$$Q(t_0) = E; \quad (71)$$

in other words, in the choice of  $n$  linearly independent solutions of (69) we shall start from the following special initial conditions:<sup>30</sup>

$$q_{ij}(t_0) = \delta_{ij} = \begin{cases} 0 & (i \neq j), \\ 1 & (i = j) \end{cases} \quad (i, j = 1, 2, \dots, n).$$

Then setting  $t = t_0$  in (70), we find from (71):

$$x_0 = c.$$

and therefore formula (70) assumes the form

$$x = Q(t)x_0 \quad (72)$$

or, in expanded form,

$$x_i = \sum_{j=1}^n q_{ij}(t) x_{j0} \quad (i=1, 2, \dots, n). \quad (72')$$

<sup>29</sup> Here the second subscript  $j$  denotes the number of the solution.

<sup>30</sup> Arbitrary initial conditions determine uniquely a certain solution of a given system.

We consider three cases:

1.  $Q(t)$  is a *bounded matrix* in the interval  $(t_0, +\infty)$ , i.e., there exists a number  $M$  such that

$$|q_{ij}(t)| \leq M \quad (t \geq t_0; i, j = 1, 2, \dots, n).$$

In this case it follows from (72') that

$$|x_i(t)| \leq nM \max |x_{j0}|.$$

The condition of stability is satisfied. (It is sufficient to take  $\delta < \frac{\varepsilon}{nM}$  in (66) and (67).) *The motion characterized by the zero solution  $x_1 = 0, x_2 = 0, \dots, x_n = 0$  is stable.*

2.  $\lim_{t \rightarrow +\infty} Q(t) = O$ . In this case the matrix  $Q(t)$  is bounded in the interval  $(t_0, +\infty)$  and therefore, as we have already explained, the motion is stable. Moreover, it follows from (72) that

$$\lim_{t \rightarrow +\infty} x(t) = o.$$

for every  $x_0$ . *The motion is asymptotically stable.*

3.  $Q(t)$  is an *unbounded matrix* in the interval  $(t_0, +\infty)$ . This means that at least one of the functions  $q_{ij}(t)$ , say  $q_{hk}(t)$ , is not bounded in the interval. We take the initial conditions  $x_{10} = 0, x_{20} = 0, \dots, x_{k-1,0} = 0, x_{k0} \neq 0, x_{k+1,0} = 0, \dots, x_{n0} = 0$ . Then

$$x_h(t) = q_{hk}(t) x_{k0}.$$

However small in modulus  $x_{k0}$  may be, the function  $x_h(t)$  is unbounded. The condition (67) is not satisfied for any  $\delta$ . *The motion is unstable.*

2. We now consider the special case where the coefficients in the system (68) are constants:

$$P(t) = P = \text{const.} \quad (73)$$

We have then (see § 5)

$$x = e^{P(t-t_0)} x_0. \quad (74)$$

Comparing (74) with (72), we find that in this case

$$Q(t) = e^{P(t-t_0)}. \quad (75)$$

We denote by

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_s)^{m_s}$$

the minimal polynomial of the coefficient matrix  $P$ .

For the investigation of the integral matrix (75) we apply formula (17) on p. 104. In this case  $f(\lambda) = e^{i(t-t_0)}$  ( $t$  is regarded as a parameter),  $f^{(i)}(\lambda_k) = (t - t_0)^i e^{i\lambda_k(t-t_0)}$ . Formula (17) yields

$$e^{P(t-t_0)} = \sum_{k=1}^s [Z_{k1} + Z_{k2}(t-t_0) + \dots + Z_{km_k}(t-t_0)^{m_k-1}] e^{i\lambda_k(t-t_0)}. \quad (78)$$

We consider three cases:

1.  $\operatorname{Re} \lambda_k \leq 0$  ( $k=1, 2, \dots, s$ ); and moreover, for all  $\lambda_k$  with  $\operatorname{Re} \lambda_k = 0$  the corresponding  $m_k = 1$  (i.e., pure imaginary characteristic values are simple roots of the minimal polynomial).
2.  $\operatorname{Re} \lambda_k < 0$  ( $k=1, 2, \dots, s$ ).
3. For some  $k$  we have  $\operatorname{Re} \lambda_k > 0$ ; or  $\operatorname{Re} \lambda_k = 0$ , but  $m_k > 1$ .

From the formula (76) it follows that in the first case the matrix  $Q(t) = e^{P(t-t_0)}$  is bounded in the interval  $(t_0, +\infty)$ , in the second case  $\lim_{t \rightarrow +\infty} e^{P(t-t_0)} = 0$ , and in the third case the matrix  $e^{P(t-t_0)}$  is not bounded in the interval  $(t_0, +\infty)$ .<sup>31</sup>

Therefore in the first case the motion  $(x_1=0, x_2=0, \dots, x_n=0)$  is stable, in the second case it is asymptotically stable, and in the third case it is unstable.

<sup>31</sup> Special consideration is only required in the case when in (76) for  $e^{P(t-t_0)}$  there occur several terms of maximal growth (for  $t \rightarrow +\infty$ ), i.e., with maximal  $\operatorname{Re} \lambda_k = \alpha_0$  and (for the given  $\operatorname{Re} \lambda_k = \alpha_0$ ) maximal value  $m_k = m_0$ . The expression (76) can be represented in the form

$$e^{P(t-t_0)} = e^{\alpha_0(t-t_0)} (t-t_0)^{m_0-1} \left[ \sum_{j=1}^r Z_{kj} m_0 e^{i\beta_j(t-t_0)} + (*) \right],$$

where  $\beta_1, \beta_2, \dots, \beta_r$  are distinct real numbers and  $(*)$  denotes a matrix that tends to zero as  $t \rightarrow +\infty$ . From this representation it follows that the matrix  $e^{P(t-t_0)}$  is not bounded for  $\alpha_0 + m_0 - 1 > 0$ , because the matrix  $\sum_{j=1}^r Z_{kj} m_0 e^{i\beta_j(t-t_0)}$  cannot converge for  $t \rightarrow +\infty$ . We can see this by showing that

$$f(t) = \sum_{j=1}^r c_j e^{i\beta_j t},$$

where  $c_j$  are complex numbers and  $\beta_j$  real and distinct numbers, can converge to zero for  $t \rightarrow +\infty$  only when  $f(t) \equiv 0$ . But, in fact, it follows from  $\lim_{t \rightarrow +\infty} f(t) = 0$  that

$$\sum_{j=1}^r |c_j|^2 = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T |f(t)|^2 dt = 0$$

and therefore

$$c_1 = c_2 = \dots = c_n = 0.$$

The results of the investigation may be formulated in the form of the following theorem.<sup>32</sup>

**THEOREM 3:** *The zero solution of the linear system (68) for  $P = \text{const.}$  is stable in the sense of Lyapunov if*

- 1) *the real parts of all the characteristic values of  $P$  are negative or zero,*
- 2) *those characteristic values whose real part is zero, i.e., the pure imaginary characteristic values (if any such exist), are simple roots of the minimal polynomial of  $P$ ;*

*and it is unstable if at least one of the conditions 1), 2) is violated.*

*The zero solution of the linear system (68) is asymptotically stable if and only if all the characteristic values of  $P$  have negative real parts.*

The considerations above enable us to make a statement about the nature of the integral matrix  $e^{P(t-t_0)}$  in the general case of arbitrary characteristic values of the constant matrix  $P$ .

**THEOREM 4:** *The integral matrix  $e^{P(t-t_0)}$  of the linear system (68) for  $P = \text{const.}$  is always representable in the form*

$$e^{P(t-t_0)} = Z_-(t) + Z_0 + Z_+(t),$$

where

- 1)  $\lim_{t \rightarrow +\infty} Z_-(t) = 0$ ,
- 2)  $Z_0$  is either constant or is a bounded matrix in the interval  $(t_0, +\infty)$  that does not have a limit for  $t \rightarrow +\infty$ ,
- 3)  $Z_+(t) = 0$  or  $Z_+(t)$  is an unbounded matrix in the interval  $(t_0, +\infty)$ .

*Proof.* On the right-hand side of (76) we divide all the summands into three groups. We denote by  $Z_-(t)$  the sum of all the terms containing the factors  $e^{\lambda_k(t-t_0)}$ , with  $\text{Re } \lambda_k < 0$ . We denote by  $Z_0$  the sum of all those matrices  $Z_{k1}$  for which  $\text{Re } \lambda_k = 0$ . We denote by  $Z_+(t)$  the sum of all the remaining terms. It is easy to see that  $Z_-(t)$ ,  $Z_0(t)$ , and  $Z_+(t)$  have the properties 1), 2), 3) of the theorem.

<sup>32</sup> On the question of sharpening the criteria of stability and instability for quasi-linear systems (i.e., of non-linear systems that become linear after neglecting the non-linear terms), see further Chapter XIV, § 3.

## CHAPTER VI

### EQUIVALENT TRANSFORMATIONS OF POLYNOMIAL MATRICES. ANALYTIC THEORY OF ELEMENTARY DIVISORS

The first three sections of this chapter deal with the theory of equivalent polynomial matrices. On the basis of this, we shall develop, in the next three sections, the analytical theory of elementary divisors, i.e., the theory of the reduction of a constant (non-polynomial) square matrix  $A$  to a normal form  $\tilde{A}$  ( $A = T\tilde{A}T^{-1}$ ). In the last two sections of the chapter two methods for the construction of the transforming matrix  $T$  will be given.

#### § 1. Elementary Transformations of a Polynomial Matrix

**1. DEFINITION 1:** A polynomial matrix, or  $\lambda$ -matrix, is a rectangular matrix  $A(\lambda)$  whose elements are polynomials in  $\lambda$ :

$$A(\lambda) = \|a_{ik}(\lambda)\| = \|a_{ik}^{(0)}\lambda^l + a_{ik}^{(1)}\lambda^{l-1} + \dots + a_{ik}^{(l)}\| \quad \begin{matrix} (i = 1, 2, \dots, m; \\ k = 1, 2, \dots, n); \end{matrix}$$

here  $l$  is the largest of the degrees of the polynomials  $a_{ik}(\lambda)$ .

Setting

$$A_j = \|a_{ik}^{(j)}\| \quad (i = 1, 2, \dots, m; k = 1, 2, \dots, n; j = 0, 1, \dots, l),$$

we may represent the polynomial matrix  $A(\lambda)$  in the form of a matrix polynomial in  $\lambda$ , i.e., in the form of a polynomial in  $\lambda$  with matrix coefficients:

$$A(\lambda) = A_0\lambda^l + A_1\lambda^{l-1} + \dots + A_{l-1}\lambda + A_l.$$

We introduce the following *elementary operations* on a polynomial matrix  $A(\lambda)$ :

1. Multiplication of any row, for example the  $i$ -th, by a number  $c \neq 0$ .

2. Addition to any row, for example the  $i$ -th, of any other row, for example the  $j$ -th, multiplied by any arbitrary polynomial  $b(\lambda)$ .

3. Interchange of any two rows, for example the  $i$ -th and the  $j$ -th.

We leave it to the reader to verify that the operations 1., 2., 3. are equivalent to a multiplication of the polynomial matrix  $A(\lambda)$  on the *left* by the following square matrices of order  $m$ , respectively:<sup>1</sup>

$$\begin{aligned}
 S' &= \left\| \begin{array}{cccccccc} & & (i) & & & & & \\ 1 & . & . & . & . & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & . & . & . & . & . & . & 1 \end{array} \right\|, \quad S'' = \left\| \begin{array}{cccccccc} & & (i) & & (j) & & & \\ 1 & . & . & . & . & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & 1 & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & . & . & . & . & . & . & 1 \end{array} \right\|, \\
 S''' &= \left\| \begin{array}{cccccccc} & & (i) & & (j) & & & \\ 1 & . & . & . & . & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & . & . & . & . & . & . & 1 \end{array} \right\|; \quad (1)
 \end{aligned}$$

in other words, as the result of applying the operations 1., 2., 3. the matrix  $A(\lambda)$  is transformed into  $S' \cdot A(\lambda)$ ,  $S'' \cdot A(\lambda)$ , and  $S''' \cdot A(\lambda)$ , respectively. The operations of type 1., 2., 3. are therefore called *left elementary operations*.

In the same way we define the *right elementary operations* on a polynomial matrix (these are performed not on the rows, but on the columns);<sup>2</sup> the matrices (of order  $n$ ) corresponding to them are:

<sup>1</sup> In the matrices (1) all the elements that are not shown are 1 on the main diagonal and 0 elsewhere.

<sup>2</sup> See footnote 1.

$$T' = \left\| \begin{array}{cccccccc} 1 & . & . & . & . & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & c & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & . & . & . & . & . & . & 1 \end{array} \right\| (i), \quad T'' = \left\| \begin{array}{cccccccc} 1 & . & . & . & . & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & 1 & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & b(\lambda) & . & . & . & . \\ . & . & . & . & . & . & . & . \\ 0 & . & . & . & . & . & . & 1 \end{array} \right\| \begin{array}{l} (i) \\ (j) \end{array}$$

$$T''' = \left\| \begin{array}{cccccccc} 1 & . & . & . & . & . & . & 0 \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & 0 & . & . & 1 & . \\ . & . & . & . & . & . & . & . \\ . & . & . & . & . & . & . & . \\ . & . & . & 1 & . & . & 0 & . \\ . & . & . & . & . & . & . & . \\ 0 & . & . & . & . & . & . & 1 \end{array} \right\| \begin{array}{l} (i) \\ (j) \end{array}$$

The result of applying a right elementary operation is equivalent to multiplying the matrix  $A(\lambda)$  on the *right* by the corresponding matrix  $T$ .

Note that  $T'$  and  $T'''$  coincide with  $S'$  and  $S'''$  and that  $T''$  coincides with  $S''$  when the indices  $i$  and  $j$  are interchanged in these matrices. The matrices of type  $S'$ ,  $S''$ ,  $S'''$  (or, what is the same,  $T'$ ,  $T''$ ,  $T'''$ ) will be called *elementary matrices*.

The determinant of every elementary matrix does not depend on  $\lambda$  and is different from zero. Therefore each left (right) elementary operation has an inverse operation which is also a left (right) elementary operation.<sup>3</sup>

**DEFINITION 2:** Two polynomial matrices  $A(\lambda)$  and  $B(\lambda)$  are called 1) *left-equivalent*, 2) *right-equivalent*, 3) *equivalent* if one of them can be obtained from the other by means of 1) *left-elementary*, 2) *right elementary*, 3) *left and right elementary operations*, respectively.<sup>4</sup>

<sup>3</sup> It follows from this that if a matrix  $B(\lambda)$  is obtained from  $A(\lambda)$  by means of left (right; left and right) elementary operations, then  $A(\lambda)$  can, conversely, be obtained from  $B(\lambda)$  by means of elementary operations of the same type. The left elementary operations form a group, as do the right elementary operations.

<sup>4</sup> From the definition it follows that only matrices of the same dimensions can be left-equivalent, right-equivalent, or simply equivalent.

Let  $B(\lambda)$  be obtained from  $A(\lambda)$  by means of the left elementary operations corresponding to  $S_1, S_2, \dots, S_p$ . Then

$$B(\lambda) = S_p S_{p-1} \cdots S_1 A(\lambda). \quad (2)$$

Denoting the product  $S_r S_{r-1} \cdots S_1$  by  $P(\lambda)$ , we write (2) in the form

$$B(\lambda) = P(\lambda) A(\lambda), \quad (3)$$

where  $P(\lambda)$ , like each of the matrices  $S_1, S_2, \dots, S_r$  has a *constant*<sup>5</sup> non-zero determinant.

In the next section we shall prove that every square  $\lambda$ -matrix  $P(\lambda)$  with a constant non-zero determinant can be represented in the form of a product of elementary matrices. Therefore (3) is equivalent to (2) and signifies left equivalence of the matrices  $A(\lambda)$  and  $B(\lambda)$ .

In the case of right equivalence of the polynomial matrices  $A(\lambda)$  and  $B(\lambda)$  we shall have instead of (3) the equation

$$B(\lambda) = A(\lambda)Q(\lambda) \quad (3')$$

and in the case of (two-sided) equivalence the equation

$$B(\lambda) = P(\lambda) A(\lambda) Q(\lambda). \quad (3'')$$

Here again,  $P(\lambda)$  and  $Q(\lambda)$  are matrices with non-zero determinants, independent of  $\lambda$ .

Thus, Definition 2 can be replaced by an equivalent definition.

DEFINITION 2': Two rectangular  $\lambda$ -matrices  $A(\lambda)$  and  $B(\lambda)$  are called

- 1) left-equivalent,
- 2) right-equivalent,
- 3) equivalent if

- 1)  $\hat{B}(\lambda) = P(\lambda)A(\lambda),$
- 2)  $B(\lambda) = A(\lambda)Q(\lambda),$
- 3)  $B(\lambda) = P(\lambda)A(\lambda)Q(\lambda).$

respectively, where  $P(\lambda)$  and  $Q(\lambda)$  are polynomial square matrices with constant non-zero determinants.

**2.** All the concepts introduced above are illustrated in the following important example.

We consider a system of  $m$  linear homogeneous differential equations of order  $l$  with constant coefficients, where  $x_1, x_2, \dots, x_n$  are  $n$  unknown functions of the independent variable  $t$ :

[illegible]

<sup>5</sup> I.e., independent of  $\lambda$ .



here

$$a_{ik}(D) = a_{ik}^{(0)}D^l + a_{ik}^{(1)}D^{l-1} + \dots + a_{ik}^{(i)} \quad (i=1, 2, \dots, m; k=1, 2, \dots, n)$$

is a polynomial in  $D$  with constant coefficients;  $D = \frac{d}{dt}$  is the differential operator.

The matrix of operator coefficients

$$A(D) = \| a_{ik}(D) \| \quad (i=1, 2, \dots, m; k=1, 2, \dots, n)$$

is a polynomial matrix, or  $D$ -matrix.

Clearly, the left elementary operation 1. on the matrix  $A(D)$  signifies term-by-term multiplication of the  $i$ -th differential equation of the system by the number  $c \neq 0$ . The left elementary operation 2. signifies the term-by-term addition to the  $i$ -th equation of the  $j$ -th equation which has previously been subjected to the differential operator  $b(D)$ . The left elementary operation 3. signifies an interchange of the  $i$ -th and  $j$ -th equation.

Thus, if we replace in (4) the matrix  $A(D)$  of operator coefficients by a left-equivalent matrix  $B(D)$ , we obtain a deduced system of equations. Since, conversely, by the same reasoning, the original system is a consequence of the new system, the two systems of equations are equivalent.<sup>6</sup>

It is not difficult in this example to interpret the right elementary operations as well. The first of them signifies the introduction of a new unknown function  $\bar{x}_i = \frac{1}{c}x_i$  for the unknown function  $x_i$ ; the second signifies the introduction of a new unknown function  $\bar{x}_j = x_j + b(D)x_i$  (instead of  $x_j$ ); the third signifies the interchange of the terms in the equations that contain  $x_i$  and  $x_j$  (i.e.,  $\bar{x}_i = x_j$ ,  $\bar{x}_j = x_i$ ).

## § 2. Canonical Form of a $\lambda$ -Matrix

1. To begin with, we shall examine what comparatively simple form we can obtain for a rectangular polynomial matrix  $A(\lambda)$  by means of left elementary operations only.

Let us assume that the first column of  $A(\lambda)$  contains elements not identically equal to zero. Among them we choose a polynomial of least degree and by a permutation of the rows we make it into the element  $a_{11}(\lambda)$ . Then we divide  $a_{i1}(\lambda)$  by  $a_{11}(\lambda)$ ; we denote quotient and remainder by  $q_{i1}(\lambda)$  and  $r_{i1}(\lambda)$  ( $i=2, \dots, m$ ):

<sup>6</sup> Here it is assumed that the unknown functions  $x_1, x_2, \dots, x_n$  are such that their derivatives of all orders, as far as they occur in the transformations, exist. With this restriction, two systems of equations with left-equivalent matrices  $A(D)$  and  $B(D)$  have the same solutions.

$$a_{i1}(\lambda) = a_{11}(\lambda)q_{i1}(\lambda) + r_{i1}(\lambda) \quad (i=2, \dots, m).$$

Now we subtract from the  $i$ -th row the first row multiplied by  $q_{i1}(\lambda)$  ( $i=2, \dots, m$ ). If not all the remainders  $r_{i1}(\lambda)$  are identically equal to zero, then we choose one of them that is not equal to zero and is of least degree and put it into the place of  $a_{11}(\lambda)$  by a permutation of the rows. As the result of all these operations, the degree of the polynomial  $a_{11}(\lambda)$  is reduced.

Now we repeat this process. Since the degree of the polynomial  $a_{11}(\lambda)$  is finite, this must come to an end at some stage—i.e., at this stage all the elements  $a_{21}(\lambda), a_{31}(\lambda), \dots, a_{m1}(\lambda)$  turn out to be identically equal to zero.

Next we take the element  $a_{22}(\lambda)$  and apply the same procedure to the rows numbered 2, 3,  $\dots, m$ , achieving  $a_{32}(\lambda) = \dots = a_{m2}(\lambda) = 0$ . Continuing still further, we finally reduce the matrix  $A(\lambda)$  to the following form:

$$\left\| \begin{array}{cccc} b_{11}(\lambda) & b_{12}(\lambda) & \dots & b_{1m}(\lambda) \\ 0 & b_{22}(\lambda) & \dots & b_{2n}(\lambda) \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{mn}(\lambda) \end{array} \right\|, \quad \left\| \begin{array}{cccc} b_{11}(\lambda) & b_{12}(\lambda) & \dots & b_{1n}(\lambda) \\ 0 & b_{22}(\lambda) & \dots & b_{2n}(\lambda) \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{nn}(\lambda) \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{array} \right\|. \quad (5)$$

$(m \leq n)$   $(m \geq n)$

If the polynomial  $b_{22}(\lambda)$  is not identically equal to zero, then by applying a left elementary operation of the second type we can make the degree of the element  $b_{12}(\lambda)$  less than the degree of  $b_{22}(\lambda)$  (if  $b_{22}(\lambda)$  is of degree zero, then  $b_{12}(\lambda)$  becomes identically equal to zero). In the same way, if  $b_{33}(\lambda) \equiv 0$ , then by left elementary operations of the second type we make the degrees of the elements  $b_{13}(\lambda), b_{23}(\lambda)$  less than the degree of  $b_{33}(\lambda)$  without changing the elements  $b_{12}(\lambda)$ , etc.

We have established the following theorem:

**THEOREM 1:** *An arbitrary rectangular polynomial matrix of dimension  $m \times n$  can always be brought into the form (5) by means of left elementary operations, where the polynomials  $b_{1k}(\lambda), b_{2k}(\lambda), \dots, b_{k-1,k}(\lambda)$  are of degree less than that of  $b_{kk}(\lambda)$ , provided  $b_{kk}(\lambda) \not\equiv 0$ , and are all identically equal to zero if  $b_{kk}(\lambda) = \text{const.} \neq 0$  ( $k=2, 3, \dots, \min(m, n)$ ).*

Similarly, we prove

**THEOREM 2:** *An arbitrary rectangular polynomial matrix of dimension  $m \times n$  can always be brought into the form*

$$\left\| \begin{array}{cccccc} c_{11}(\lambda) & 0 & \dots & 0 & 0 & \dots & 0 \\ c_{21}(\lambda) & c_{22}(\lambda) & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ c_{m1}(\lambda) & c_{m2}(\lambda) & \dots & c_{mm}(\lambda) & 0 & \dots & 0 \end{array} \right\|, \quad \left\| \begin{array}{cccc} c_{11}(\lambda) & 0 & \dots & 0 \\ c_{21}(\lambda) & c_{22}(\lambda) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ c_{n1}(\lambda) & c_{n2}(\lambda) & \dots & c_{nn}(\lambda) \\ \dots & \dots & \dots & \dots \\ c_{m1}(\lambda) & c_{m2}(\lambda) & \dots & c_{mn}(\lambda) \end{array} \right\| \quad (6)$$

$(m \leq n)$   $(n \geq m)$

by means of right elementary operations, where the polynomials  $c_{k1}(\lambda)$ ,  $c_{k2}(\lambda)$ , ...,  $c_{k, k-1}(\lambda)$  are of degree less than that of  $c_{kk}(\lambda)$ , provided  $c_{kk}(\lambda) \neq 0$ , and all are identically equal to zero if  $c_{kk}(\lambda) = \text{const.} \neq 0$  ( $k = 2, 3, \dots, \min(m, n)$ ).

2. From Theorems 1 and 2 we deduce the corollary:

**COROLLARY:** If the determinant of a square polynomial matrix  $P(\lambda)$  does not depend on  $\lambda$  and is different from zero, then the matrix can be represented in the form of a product of a finite number of elementary matrices.

For by Theorem 1 the matrix  $P(\lambda)$  can be brought into the form

$$\left\| \begin{array}{cccc} b_{11}(\lambda) & b_{12}(\lambda) & \dots & b_{1n}(\lambda) \\ 0 & b_{22}(\lambda) & \dots & b_{2n}(\lambda) \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{nn}(\lambda) \end{array} \right\| \quad (7)$$

by left elementary operations, where  $n$  is the order of  $P(\lambda)$ . Since in the application of elementary operations to a square polynomial matrix the determinant of the matrix is only multiplied by constant non-zero factors, the determinant of the matrix (7), like that of  $P(\lambda)$ , does not depend on  $\lambda$  and is different from 0, i.e.,

$$b_{11}(\lambda) b_{22}(\lambda) \dots b_{nn}(\lambda) = \text{const.} \neq 0.$$

Hence

$$b_{kk}(\lambda) = \text{const.} \neq 0 \quad (k = 1, 2, \dots, n).$$

But then, also by Theorem 1, the matrix (7) has the diagonal form  $\|b_k \delta_{ik}\|_1^n$  and can therefore be reduced to the unit matrix  $E$  by means of left elementary operations of type 1. But then, conversely, the unit matrix  $E$  can be transformed into  $P(\lambda)$  by means of the left elementary operations whose matrices are  $S_1, S_2, \dots, S_p$ . Therefore

$$P(\lambda) = S_p S_{p-1} \dots S_1 E = S_p S_{p-1} \dots S_1.$$



If at least one of the elements  $a_{ik}(\lambda)$  ( $i = 2, \dots, m; k = 2, \dots, n$ ) is not divisible without remainder by  $a_{11}(\lambda)$ , then by adding to the first column that column which contains such an element we arrive at the preceding case and can therefore again replace the element  $a_{11}(\lambda)$  by a polynomial of smaller degree.

Since the original element  $a_{11}(\lambda)$  had a definite degree and since the process of reducing this degree cannot be continued indefinitely, we must, after a finite number of elementary operations, obtain a matrix of the form

$$\left\| \begin{array}{cccc} a_1(\lambda) & 0 & \dots & 0 \\ 0 & b_{22}(\lambda) & \dots & b_{2n}(\lambda) \\ \dots & \dots & \dots & \dots \\ 0 & b_{m2}(\lambda) & \dots & b_{mn}(\lambda) \end{array} \right\|, \quad (8)$$

in which all the elements  $b_{ik}(\lambda)$  are divisible without remainder by  $a_1(\lambda)$ . If among these elements  $b_{ik}(\lambda)$  there is one not identically equal to zero, then continuing the same reduction process on the rows numbered  $2, \dots, m$  and the columns  $2, \dots, n$ , we reduce the matrix (8) to the form

$$\left\| \begin{array}{cccccc} a_1(\lambda) & 0 & 0 & \dots & 0 \\ 0 & a_2(\lambda) & 0 & \dots & 0 \\ 0 & 0 & c_{33}(\lambda) & \dots & c_{3n}(\lambda) \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & c_{m3}(\lambda) & \dots & c_{mn}(\lambda) \end{array} \right\|,$$

where  $a_2(\lambda)$  is divisible without remainder by  $a_1(\lambda)$  and all the polynomials  $c_{ik}(\lambda)$  are divisible without remainder by  $a_2(\lambda)$ . Continuing the process further, we finally arrive at a matrix of the form

$$\left\| \begin{array}{cccccc} a_1(\lambda) & 0 & \dots & 0 & 0 \dots 0 \\ 0 & a_2(\lambda) & \dots & 0 & 0 \dots 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_s(\lambda) & 0 \dots 0 \\ 0 & 0 & \dots & 0 & 0 \dots 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 \dots 0 \end{array} \right\|, \quad (9)$$

where the polynomials  $a_1(\lambda), a_2(\lambda), \dots, a_s(\lambda)$  ( $s \leq \min(m, n)$ ) are not identically equal to zero and each is divisible by the preceding one.

By multiplying the first  $s$  rows by suitable non-zero numerical factors, we can arrange that the highest coefficients of the polynomials  $a_1(\lambda), a_2(\lambda), \dots, a_s(\lambda)$  are equal to 1.

DEFINITION 3: A rectangular polynomial matrix is called a canonical diagonal matrix if it is of the form (9), where 1) the polynomials  $a_1(\lambda)$ ,  $a_2(\lambda)$ ,  $\dots$ ,  $a_s(\lambda)$  are not identically equal to zero and 2) each of the polynomials  $a_2(\lambda)$ ,  $\dots$ ,  $a_s(\lambda)$  is divisible by the preceding. Moreover, it is assumed that the highest coefficients of all the polynomials  $a_1(\lambda)$ ,  $a_2(\lambda)$ ,  $\dots$ ,  $a_s(\lambda)$  are equal to 1.

Thus, we have proved that: An arbitrary rectangular polynomial matrix  $A(\lambda)$  is equivalent to a canonical diagonal matrix. In the next section we shall prove that: The polynomials  $a_1(\lambda)$ ,  $a_2(\lambda)$ ,  $\dots$ ,  $a_s(\lambda)$  are uniquely determined by the given matrix  $A(\lambda)$ ; and we shall set up formulas that connect these polynomials with the elements of  $A(\lambda)$ .

### § 3. Invariant Polynomials and Elementary Divisors of a Polynomial Matrix

1. We introduce the concept of invariant polynomials of a  $\lambda$ -matrix  $A(\lambda)$ .

Let  $A(\lambda)$  be a polynomial matrix of rank  $r$ , i.e., the matrix has minors of order  $r$  not identically equal to zero, but all the minors of order greater than  $r$  are identically equal to zero in  $\lambda$ . We denote by  $D_j(\lambda)$  the greatest common divisor of all the minors of order  $j$  in  $A(\lambda)$  ( $j = 1, 2, \dots, r$ ).<sup>7</sup> Then it is easy to see that in the series

$$D_r(\lambda), D_{r-1}(\lambda), \dots, D_1(\lambda), D_0(\lambda) \equiv 1$$

each polynomial is divisible by the preceding one.<sup>8</sup> The corresponding quotients will be denoted by  $i_1(\lambda)$ ,  $i_2(\lambda)$ ,  $\dots$ ,  $i_r(\lambda)$ :

$$i_1(\lambda) = \frac{D_r(\lambda)}{D_{r-1}(\lambda)}, \quad i_2(\lambda) = \frac{D_{r-1}(\lambda)}{D_{r-2}(\lambda)}, \quad \dots, \quad i_r(\lambda) = \frac{D_1(\lambda)}{D_0(\lambda)} = D_1(\lambda). \quad (10)$$

DEFINITION 4: The polynomials  $i_1(\lambda)$ ,  $i_2(\lambda)$ ,  $\dots$ ,  $i_r(\lambda)$  defined by (10) are called the invariant polynomials of the rectangular matrix  $A(\lambda)$ .

The term 'invariant polynomial' is explained by the following arguments. Let  $A(\lambda)$  and  $B(\lambda)$  be two equivalent polynomial matrices. Then they are obtained from one another by means of elementary operations. But an easy verification shows immediately that the elementary operations

<sup>7</sup> We take the highest coefficient in  $D_j(\lambda)$  to be 1 ( $j = 1, 2, \dots, r$ ).

<sup>8</sup> If we apply the Bézout decomposition with respect to the elements of any row to an arbitrary minor of order  $j$ , then every term in the decomposition is divisible by  $D_{j-1}(\lambda)$ ; therefore every minor of order  $j$ , and hence  $D_j(\lambda)$ , is divisible by  $D_{j-1}(\lambda)$  ( $j = 2, 3, \dots, r$ ).

change neither the rank of  $A(\lambda)$  nor the polynomials  $D_1(\lambda)$ ,  $D_2(\lambda)$ , ...,  $D_r(\lambda)$ . For when we apply to the identity (3'') the formula that expresses a minor of a product of matrices by the minors of the factors (see p 12), we obtain for an arbitrary minor of  $B(\lambda)$  the expression

$$B \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} = \sum_{\substack{1 \leq \alpha_1 < \alpha_2 < \dots < \alpha_p \leq m \\ 1 \leq \beta_1 < \beta_2 < \dots < \beta_p \leq m}} P \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ \alpha_1 & \alpha_2 & \dots & \alpha_p \end{pmatrix} A \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_p \\ \beta_1 & \beta_2 & \dots & \beta_p \end{pmatrix} Q \begin{pmatrix} \beta_1 & \beta_2 & \dots & \beta_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \\ (p=1, 2, \dots, \min(m, n)).$$

Hence it follows that all the minors of order  $r$  or greater of the matrix  $B(\lambda)$  are zero, so that we have for the rank  $r^*$  of  $B(\lambda)$ :

$$r^* \leq r.$$

Moreover, it follows from the same formula that  $D_p^*(\lambda)$ , the greatest common divisor of all the minors of order  $p$  of  $B(\lambda)$ , is divisible by  $D_p(\lambda)$  ( $p=1, 2, \dots, \min(m, n)$ ). But the matrices  $A(\lambda)$  and  $B(\lambda)$  can exchange roles. Therefore  $r \leq r^*$  and  $D_p(\lambda)$  is divisible by  $D_p^*(\lambda)$  ( $p=1, 2, \dots, \min(m, n)$ ). Hence<sup>9</sup>

$$r=r^*, \quad D_1^*(\lambda)=D_1(\lambda), \quad D_2^*(\lambda)=D_2(\lambda), \dots, \quad D_r^*(\lambda)=D_r(\lambda).$$

Since elementary operations do not change the polynomials  $D_1(\lambda)$ ,  $D_2(\lambda)$ , ...,  $D_r(\lambda)$ , they also leave the polynomials  $i_1(\lambda)$ ,  $i_2(\lambda)$ , ...,  $i_r(\lambda)$  defined by (10) unchanged.

Thus, the polynomials  $i_1(\lambda)$ ,  $i_2(\lambda)$ , ...,  $i_r(\lambda)$  remain *invariant* on transition from one matrix to another equivalent one.

If the polynomial matrix has the canonical diagonal form (9), then it is easy to see that for this matrix

$$D_1(\lambda)=a_1(\lambda), \quad D_2(\lambda)=a_1(\lambda)a_2(\lambda), \quad \dots, \quad D_r(\lambda)=a_1(\lambda)a_2(\lambda)\dots a_r(\lambda).$$

But then, by (10), the diagonal polynomials in (9)  $a_1(\lambda)$ ,  $a_2(\lambda)$ , ...  $a_r(\lambda)$  coincide with the invariant polynomials

$$i_1(\lambda)=a_r(\lambda), \quad i_2(\lambda)=a_{r-1}(\lambda), \quad \dots, \quad i_r(\lambda)=a_1(\lambda). \quad (11)$$

Here  $i_1(\lambda)$ ,  $i_2(\lambda)$ , ...,  $i_r(\lambda)$  are at the same time the invariant polynomials of the original matrix  $A(\lambda)$ , because it is equivalent to (9).

The results obtained can be stated in the form of the following theorem.

<sup>9</sup> The highest coefficients in  $D_p(\lambda)$  and  $D_p^*(\lambda)$  ( $p=1, 2, \dots, r$ ) are 1.

**THEOREM 3:** *The rectangular polynomial matrix  $A(\lambda)$  is always equivalent to a canonical diagonal matrix*

$$\left\| \begin{array}{cccccc} i_r(\lambda) & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & i_{r-1}(\lambda) & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & i_1(\lambda) & 0 & \dots & 0 \\ 0 & 0 & & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & & 0 & 0 & \dots & 0 \end{array} \right\|. \quad (12)$$

Moreover,  $r$  must here be the rank of  $A(\lambda)$  and  $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$  the invariant polynomials of  $A(\lambda)$  defined by (10).

**COROLLARY 1:** *Two rectangular matrices of the same dimension  $A(\lambda)$  and  $B(\lambda)$  are equivalent if and only if they have the same invariant polynomials.*

The sufficiency of the condition was explained above. The necessity follows from the fact that two polynomial matrices having the same invariant polynomials are equivalent to one and the same canonical diagonal matrix and, therefore, to each other. Thus: *The invariant polynomials form a complete system of invariants of a  $\lambda$ -matrix.*

**COROLLARY 2:** *In the sequence of invariant polynomials*

$$i_1(\lambda) = \frac{D_r(\lambda)}{D_{r-1}(\lambda)}, \quad i_2(\lambda) = \frac{D_{r-1}(\lambda)}{D_{r-2}(\lambda)}, \quad \dots, \quad i_r(\lambda) = \frac{D_1(\lambda)}{D_0(\lambda)} \quad (D_0(\lambda) \equiv 1) \quad (13)$$

*every polynomial from the second onwards divides the preceding one.*

This statement does not follow immediately from (13). It does follow from the fact that the polynomials  $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$  coincide with the polynomials  $a_r(\lambda), a_{r-1}(\lambda), \dots, a_1(\lambda)$  of the canonical diagonal matrix (9).

**2.** We now indicate a method of computing the invariant polynomials of a quasi-diagonal  $\lambda$ -matrix if the invariant polynomials of the matrices in the diagonal blocks are known.

**THEOREM 4:** *If in a quasi-diagonal rectangular matrix*

$$C(\lambda) = \left\| \begin{array}{cc} A(\lambda) & O \\ O & B(\lambda) \end{array} \right\|$$

*every invariant polynomial of  $A(\lambda)$  divides every invariant polynomial of  $B(\lambda)$ , then the set of invariant polynomials of  $C(\lambda)$  is the union of the invariant polynomials of  $A(\lambda)$  and  $B(\lambda)$ .*



*Proof.* We denote by  $i'_1(\lambda), i'_2(\lambda), \dots, i'_r(\lambda)$  and  $i''_1(\lambda), i''_2(\lambda), \dots, i''_q(\lambda)$ , respectively, the invariant polynomials of the  $\lambda$ -matrices  $A(\lambda)$  and  $B(\lambda)$ . Then<sup>10</sup>

$$A(\lambda) \sim \{i'_1(\lambda), \dots, i'_r(\lambda), 0, \dots, 0\}, \quad B(\lambda) \sim \{i''_1(\lambda), \dots, i''_q(\lambda), 0, \dots, 0\}$$

and therefore

$$C(\lambda) \sim \{i'_r(\lambda), \dots, i'_1(\lambda), i''_q(\lambda), \dots, i''_1(\lambda), 0, \dots, 0\}. \quad (14)$$

The  $\lambda$ -matrix on the right-hand side of this relation is of canonical diagonal form. By Theorem 3 the diagonal elements of this matrix that are not identically equal to zero then form a complete system of invariants of the polynomial matrix  $C(\lambda)$ . This proves the theorem.

In order to determine the invariant polynomials of  $C(\lambda)$  in the general case of arbitrary invariant polynomials of  $A(\lambda)$  and  $B(\lambda)$  we make use of the important concept of elementary divisors.

We decompose the invariant polynomials  $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$  into irreducible factors over the given number field  $F$ :<sup>11</sup>

$$\begin{aligned} i_1(\lambda) &= [\varphi_1(\lambda)]^{c_1} [\varphi_2(\lambda)]^{c_2} \dots [\varphi_s(\lambda)]^{c_s}, \\ i_2(\lambda) &= [\varphi_1(\lambda)]^{d_1} [\varphi_2(\lambda)]^{d_2} \dots [\varphi_s(\lambda)]^{d_s}, \\ &\dots \dots \dots \dots \dots \dots \dots \left( \begin{array}{l} c_k \geq d_k \geq \dots \geq l_k \geq 0; \\ k = 1, 2, \dots, s \end{array} \right) \\ i_r(\lambda) &= [\varphi_1(\lambda)]^{l_1} [\varphi_2(\lambda)]^{l_2} \dots [\varphi_s(\lambda)]^{l_s}. \end{aligned} \quad (15)$$

Here  $\varphi_1(\lambda), \varphi_2(\lambda), \dots, \varphi_s(\lambda)$  are all the distinct factors irreducible over  $F$  (and with highest coefficient 1) that occur in  $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$ .

**DEFINITION 5:** All the powers among  $[\varphi_1(\lambda)]^{c_1}, \dots, [\varphi_s(\lambda)]^{l_s}$  in (15), as far as they are distinct from 1, are called the elementary divisors of the matrix  $A(\lambda)$  in the field  $F$ .<sup>12</sup>

**THEOREM 5:** The set of elementary divisors of the rectangular quasi-diagonal matrix

$$C(\lambda) = \left\| \begin{array}{cc} A(\lambda) & O \\ O & B(\lambda) \end{array} \right\|$$

is always obtained by combining the elementary divisors of  $A(\lambda)$  with those of  $B(\lambda)$ .

<sup>10</sup> The symbol  $\sim$  denotes here the equivalence of matrices; and braces  $\{ \}$ , a diagonal rectangular matrix of the form (12).

<sup>11</sup> Some of the exponents  $c_k, d_k, \dots, l_k$  ( $k = 1, 2, \dots, s$ ) may be equal to zero.

<sup>12</sup> The formulas (15) enable us to define not only the elementary divisors of  $A(\lambda)$  in the field  $F$  in terms of the invariant polynomials but also, conversely, the invariant polynomials in terms of the elementary divisors.

*Proof.* We decompose the invariant polynomials of  $A(\lambda)$  and  $B(\lambda)$  into irreducible factors over  $\mathbb{F}$ :<sup>13</sup>

$$\begin{aligned} i'_1(\lambda) &= [\varphi_1(\lambda)]^{c'_1} [\varphi_2(\lambda)]^{d'_1} \dots [\varphi_s(\lambda)]^{g'_1}, & i''_1(\lambda) &= [\varphi_1(\lambda)]^{c''_1} [\varphi_2(\lambda)]^{d''_1} \dots [\varphi_s(\lambda)]^{g''_1}, \\ i'_2(\lambda) &= [\varphi_1(\lambda)]^{c'_2} [\varphi_2(\lambda)]^{d'_2} \dots [\varphi_s(\lambda)]^{g'_2}, & i''_2(\lambda) &= [\varphi_1(\lambda)]^{c''_2} [\varphi_2(\lambda)]^{d''_2} \dots [\varphi_s(\lambda)]^{g''_2}, \\ &\dots & & \dots \\ i'_r(\lambda) &= [\varphi_1(\lambda)]^{c'_r} [\varphi_2(\lambda)]^{d'_r} \dots [\varphi_s(\lambda)]^{g'_r}, & i''_r(\lambda) &= [\varphi_1(\lambda)]^{c''_r} [\varphi_2(\lambda)]^{d''_r} \dots [\varphi_s(\lambda)]^{g''_r}. \end{aligned}$$

We denote by

$$c_1 \geq d_1 \geq \dots \geq l_1 > 0, \quad (16)$$

all the non-zero numbers among  $c'_1, d'_1, \dots, h'_1, c''_1, d''_1, \dots, g''_1$ .

Then the matrix  $C(\lambda)$  is equivalent to the matrix (14), and by a permutation of rows and of columns the latter can be brought into 'diagonal' form

$$\{[\varphi_1(\lambda)]^{c_1} \cdot (*), [\varphi_1(\lambda)]^{d_1} \cdot (*), \dots, [\varphi_1(\lambda)]^{l_1} \cdot (*), (**), \dots, (**)\} \quad (17)$$

where we have denoted by  $(*)$  polynomials that are prime to  $\varphi_1(\lambda)$  and by  $(**)$  polynomials that are either prime to  $\varphi_1(\lambda)$  or identically equal to zero. From the form of the matrix (17) we deduce immediately the following decomposition of the polynomials  $D_r(\lambda)$ ,  $D_{r-1}(\lambda)$ ,  $\dots$  and  $i_1(\lambda)$ ,  $i_2(\lambda)$ ,  $\dots$  of the matrix  $C(\lambda)$ :

$$\begin{aligned} D_r(\lambda) &= [\varphi_1(\lambda)]^{c_1+d_1+\dots+l_1} \cdot (*), & D_{r-1}(\lambda) &= [\varphi_1(\lambda)]^{d_1+\dots+l_1} \cdot (*), \dots, \\ i_1(\lambda) &= [\varphi_1(\lambda)]^{c_1} (*), & i_2(\lambda) &= [\varphi_1(\lambda)]^{d_1} (*), \dots \end{aligned}$$

Hence it follows that  $[\varphi_1(\lambda)]^{c_1}$ ,  $[\varphi_1(\lambda)]^{d_1}$ ,  $\dots$ ,  $[\varphi_1(\lambda)]^{l_1}$ , i.e., all the powers

$$[\varphi_1(\lambda)]^{c'_1}, \dots, [\varphi_1(\lambda)]^{h'_1}, [\varphi_1(\lambda)]^{c''_1}, \dots, [\varphi_1(\lambda)]^{g''_1},$$

as far as they are distinct from 1, are elementary divisors of  $C(\lambda)$ .

The elementary divisors of  $C(\lambda)$  that are powers of  $\varphi_2(\lambda)$  are determined similarly, etc. This completes the proof of the theorem.

*Note.* The theory of equivalence for integral matrices (i.e., matrices whose elements are integers) can be constructed along similar lines. Here in 1., 2. (see pp. 130-31)  $c = \pm 1$ ,  $b(\lambda)$  is to be replaced by an integer, and in (3), (3'), (3''), in place of  $P(\lambda)$  and  $Q(\lambda)$  there are integral matrices with determinants equal to  $\pm 1$ .

<sup>13</sup> If any irreducible polynomial  $\varphi_k(\lambda)$  occurs as a factor in some invariant polynomials, but not in others, then in the latter we write  $\varphi_k(\lambda)$  with a zero exponent.

3. Suppose now that  $A = \|a_{ik}\|_1^n$  is a matrix with elements in the field  $\mathbb{F}$ . We form its characteristic matrix

$$\lambda E - A = \left\| \begin{array}{cccc} \lambda - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ -a_{n1} & -a_{n2} & \dots & \lambda - a_{nn} \end{array} \right\|. \quad (18)$$

The characteristic matrix is a  $\lambda$ -matrix of rank  $n$ . Its invariant polynomials

$$i_1(\lambda) = \frac{D_n(\lambda)}{D_{n-1}(\lambda)}, \quad i_2(\lambda) = \frac{D_{n-1}(\lambda)}{D_{n-2}(\lambda)}, \quad \dots, \quad i_n(\lambda) = \frac{D_1(\lambda)}{D_0(\lambda)} \quad (D_0(\lambda) \equiv 1), \quad (19)$$

are called the *invariant polynomials of the matrix A* and the corresponding elementary divisors in  $\mathbb{F}$  are called the *elementary divisors of the matrix A in the field F*. A knowledge of the invariant polynomials (and, hence, of the elementary divisors) of  $A$  enables us to investigate its structure. Therefore practical methods of computing the invariant polynomials of a matrix are of interest. The formulas (19) give an algorithm for computing these polynomials, but for large  $n$  this algorithm is very cumbersome.

Theorem 3 gives another method of computing invariant polynomials, based on the reduction of the characteristic matrix (18) to canonical diagonal form by means of elementary operations.

*Example:*

$$A = \left\| \begin{array}{cccc} 3 & 1 & 0 & 0 \\ -4 & -1 & 0 & 0 \\ 6 & 1 & 2 & 1 \\ -14 & -5 & -1 & 0 \end{array} \right\|, \quad \lambda E - A = \left\| \begin{array}{cccc} \lambda - 3 & -1 & 0 & 0 \\ 4 & \lambda + 1 & 0 & 0 \\ -6 & -1 & \lambda - 2 & -1 \\ 14 & 5 & 1 & \lambda \end{array} \right\|.$$

In the characteristic matrix  $\lambda E - A$  we add to the fourth row the third multiplied by  $\lambda$ :

$$\left\| \begin{array}{cccc} \lambda - 3 & -1 & 0 & 0 \\ 4 & \lambda + 1 & 0 & 0 \\ -6 & -1 & \lambda - 2 & -1 \\ 14 - 6\lambda & 5 - \lambda & \lambda^2 - 2\lambda + 1 & 0 \end{array} \right\|.$$

Now adding to the first three columns the fourth, multiplied by  $-6$ ,  $-1$  and  $\lambda - 2$ , respectively, we obtain

$$\left\| \begin{array}{cccc} \lambda - 3 & -1 & 0 & 0 \\ 4 & \lambda + 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 14 - 6\lambda & 5 - \lambda & \lambda^2 - 2\lambda + 1 & 0 \end{array} \right\|.$$

We add to the first column the second multiplied by  $\lambda - 3$ :

$$\begin{vmatrix} 0 & -1 & 0 & 0 \\ \lambda^2 - 2\lambda + 1 & \lambda + 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ -\lambda^2 + 2\lambda - 1 & 5 - \lambda & \lambda^2 - 2\lambda + 1 & 0 \end{vmatrix}.$$

To the second and fourth rows we add the first multiplied by  $\lambda + 1$  and  $5 - \lambda$ , respectively; we obtain

$$\begin{vmatrix} 0 & -1 & 0 & 0 \\ \lambda^2 - 2\lambda + 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ -\lambda^2 + 2\lambda - 1 & 0 & \lambda^2 - 2\lambda + 1 & 0 \end{vmatrix}.$$

To the second row we add the fourth; then we multiply the first and third rows by  $-1$ . After permuting some rows and columns we obtain:

$$\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & (\lambda - 1)^2 & 0 \\ 0 & 0 & 0 & (\lambda - 1)^2 \end{vmatrix}.$$

The matrix has two elementary divisors  $(\lambda - 1)^2$  and  $(\lambda - 1)^2$ .

#### § 4. Equivalence of Linear Binomials

1. In the preceding sections we have considered rectangular  $\lambda$ -matrices. In the present section we consider two square  $\lambda$ -matrices  $A(\lambda)$  and  $B(\lambda)$  of order  $n$  in which all the elements are of degree not higher than 1 in  $\lambda$ . These polynomial matrices may be represented in the form of matrix binomials:

$$A(\lambda) = A_0\lambda + A_1, \quad B(\lambda) = B_0\lambda + B_1.$$

We shall assume that these binomials are of degree 1 and regular, i.e., that  $|A_0| \neq 0$ ,  $|B_0| \neq 0$  (see p. 76).

The following theorem gives a criterion for the equivalence of such binomials:

**THEOREM 6:** *If two regular binomials of the first degree  $A_0\lambda + A_1$  and  $B_0\lambda + B_1$  are equivalent, then they are strictly equivalent, i.e., in the identity*

$$B_0\lambda + B_1 = P(\lambda)(A_0\lambda + A_1)Q(\lambda) \quad (20)$$

*the matrices  $P(\lambda)$  and  $Q(\lambda)$ —with constant non-zero determinants—can be replaced by constant non-singular matrices  $P$  and  $Q$ :<sup>14</sup>*

$$B_0\lambda + B_1 = P(A_0\lambda + A_1)Q. \quad (21)$$

<sup>14</sup> The identity (21) is equivalent to the two matrix equations:  $B_0 = PA_0Q$  and  $B_1 = PA_1Q$ .

*Proof.* Since the determinant of  $P(\lambda)$  does not depend on  $\lambda$  and is different from zero,<sup>15</sup> the inverse matrix  $M(\lambda) = P^{-1}(\lambda)$  is also a polynomial matrix. With the help of this matrix we write (20) in the form

$$M(\lambda)(B_0\lambda + B_1) = (A_0\lambda + A_1)Q(\lambda). \quad (22)$$

Regarding  $M(\lambda)$  and  $Q(\lambda)$  as matrix polynomials, we divide  $M(\lambda)$  on the left by  $A_0\lambda + A_1$  and  $Q(\lambda)$  on the right by  $B_0\lambda + B_1$ :

$$M(\lambda) = (A_0\lambda + A_1)S(\lambda) + M, \quad (23)$$

$$Q(\lambda) = T(\lambda)(B_0\lambda + B_1) + Q; \quad (24)$$

here  $M$  and  $Q$  are constant square matrices (independent of  $\lambda$ ) of order  $n$ . We substitute these expressions for  $M(\lambda)$  and  $Q(\lambda)$  in (22). After a few small transformations, we obtain

$$(A_0\lambda + A_1)[T(\lambda) - S(\lambda)](B_0\lambda + B_1) = M(B_0\lambda + B_1) - (A_0\lambda + A_1)Q. \quad (25)$$

The difference in the brackets must be identically equal to zero; for otherwise the product on the left-hand side of (25) would be of degree  $\geq 2$ , while the polynomial on the right-hand side of the equation is of degree not higher than 1. Therefore

$$S(\lambda) = T(\lambda); \quad (26)$$

But then we obtain from (25):

$$M(B_0\lambda + B_1) = (A_0\lambda + A_1)Q. \quad (27)$$

We shall now show that  $M$  is a non-singular matrix. For this purpose we divide  $P(\lambda)$  on the left by  $B_0\lambda + B_1$ :

$$P(\lambda) = (B_0\lambda + B_1)U(\lambda) + P. \quad (28)$$

From (22), (23), and (28) we deduce:

$$\begin{aligned} E &= M(\lambda)P(\lambda) = M(\lambda)(B_0\lambda + B_1)U(\lambda) + M(\lambda)P \\ &= (A_0\lambda + A_1)Q(\lambda)U(\lambda) + (A_0\lambda + A_1)S(\lambda)P + MP \\ &= (A_0\lambda + A_1)[Q(\lambda)U(\lambda) + S(\lambda)P] + MP. \end{aligned} \quad (29)$$

<sup>15</sup> The equivalence of the binomials  $A_0\lambda + A_1$  and  $B_0\lambda + B_1$  means that an identity (20) exists in which  $|P(\lambda)| = \text{const.} \neq 0$  and  $|Q(\lambda)| = \text{const.} \neq 0$ . However, in this case the last relations follow from (20) itself. For the determinants of regular binomials of the first degree are of degree  $n$ :

$$|A_0\lambda + A_1| = |A_0|\lambda^n + \dots, |B_0\lambda + B_1| = |B_0|\lambda^n + \dots; |A_0| \neq 0, |B_0| \neq 0.$$

Therefore it follows from

$$|B_0\lambda + B_1| = |P(\lambda)| |A_0\lambda + A_1| |Q(\lambda)|$$

that

$$|P(\lambda)| = \text{const.} \neq 0, |Q(\lambda)| = \text{const.} \neq 0.$$

Since the last term of this chain of equations must be of degree zero in  $\lambda$  (because it is equal to  $E$ ), the expression in brackets must be identically equal to zero. But then from (29)

$$MP = E, \quad (30)$$

so that  $|M| \neq 0$  and  $M^{-1} = P$ .

Multiplying both sides of (27) on the left by  $P$ , we obtain:

$$B_0\lambda + B_1 = P(A_0\lambda + A_1)Q.$$

The fact that  $P$  is non-singular follows from (30). That  $P$  and  $Q$  are non-singular also follows directly from (21), since this identity implies

$$B_0 = PA_0Q$$

and therefore

$$|P||A_0||Q| = |B_0| \neq 0.$$

This completes the proof of the theorem.

*Note.* From the proof it follows (see (24) and (28)) that the constant matrices  $P$  and  $Q$  by which we have replaced the  $\lambda$ -matrices  $P(\lambda)$  and  $Q(\lambda)$  in (20) can be taken as the left and right remainders, respectively, of  $P(\lambda)$  and  $Q(\lambda)$  on division by  $B_0\lambda + B_1$ .

## § 5. A Criterion for Similarity of Matrices

1. Let  $A = \|a_{ik}\|_1^n$  be a matrix with numerical elements from the field  $\mathbb{F}$ . Its characteristic matrix  $\lambda E - A$  is a  $\lambda$ -matrix of rank  $n$  and therefore has  $n$  invariant polynomials (see § 3)

$$i_1(\lambda), i_2(\lambda), \dots, i_n(\lambda).$$

The following theorem shows that these invariant polynomials determine the original matrix  $A$  to within similarity transformations.

**THEOREM 7:** *Two matrices  $A = \|a_{ik}\|_1^n$  and  $B = \|b_{ik}\|_1^n$  are similar ( $B = T^{-1}AT$ ) if and only if they have the same invariant polynomials or, what is the same, the same elementary divisors in the field  $\mathbb{F}$ .*

*Proof.* The condition is *necessary*. For if the matrices  $A$  and  $B$  are similar, then there exists a non-singular matrix  $T$  such that

$$B = T^{-1}AT.$$

Hence

$$\lambda E - B = T^{-1}(\lambda E - A)T.$$

This equation shows that the characteristic matrices  $\lambda E - A$  and  $\lambda E - B$  are equivalent and therefore have the same invariant polynomials.

The condition is *sufficient*. Suppose that the characteristic matrices  $\lambda E - A$  and  $\lambda E - B$  have the same invariant polynomials. Then these  $\lambda$ -matrices are equivalent (see Corollary 1 to Theorem 3) and there exist, in consequence, two polynomial matrices  $P(\lambda)$  and  $Q(\lambda)$  such that

$$\lambda E - B = P(\lambda)(\lambda E - A)Q(\lambda). \quad (31)$$

Applying Theorem 6 to the matrix binomials  $\lambda E - A$  and  $\lambda E - B$ , we may replace in (31) the  $\lambda$ -matrices  $P(\lambda)$  and  $Q(\lambda)$  by constant matrices  $P$  and  $Q$ :

$$\lambda E - B = P(\lambda E - A)Q; \quad (32)$$

moreover,  $P$  and  $Q$  may be taken (see the Note on p. 147) as the left remainder and the right remainder, respectively, of  $P(\lambda)$  and  $Q(\lambda)$  on division by  $\lambda E - B$ , i.e., by the Generalized Bézout Theorem, we may set:<sup>16</sup>

$$P = \hat{P}(B), \quad Q = Q(B) \quad (33)$$

Equating coefficients of the powers of  $\lambda$  on both sides of (32), we obtain:

$$B = PAQ, \quad E = PQ,$$

i.e.,

$$B = T^{-1}AT,$$

where

$$T = Q = P^{-1}$$

This proves the theorem.

**2. Note.** We have incidentally established the following result, which we state separately:

**SUPPLEMENT TO THEOREM 7.** If  $A = \|a_{ik}\|_1^n$  and  $B = \|b_{ik}\|_1^n$  are two similar matrices,

$$B = T^{-1}AT, \quad (34)$$

then we can choose as the transforming matrix  $T$  the matrix

$$T = Q(B) = [\hat{P}(B)]^{-1}, \quad (35)$$

where  $P(\lambda)$  and  $Q(\lambda)$  are polynomial matrices in the identity

$$\lambda E - B = P(\lambda)(\lambda E - A)Q(\lambda)$$

which connects the equivalent characteristic matrices  $\lambda E - A$  and  $\lambda E - B$ ; in (35)  $Q(B)$  denotes the right value of the matrix polynomial  $Q(\lambda)$ , and  $\hat{P}(B)$  the left value of  $P(\lambda)$ , when the argument is replaced by  $B$ .

<sup>16</sup> We recall that  $\hat{P}(B)$  is the left value of the polynomial  $P(\lambda)$  and  $Q(B)$  the right value of  $Q(\lambda)$ , when  $\lambda$  is replaced by  $B$  (see p. 81).

## § 6. The Normal Forms of a Matrix

1. Let

$$g(\lambda) = \lambda^m + \alpha_1 \lambda^{m-1} + \dots + \alpha_{m-1} \lambda + \alpha_m$$

be a polynomial with coefficients in  $\mathbb{F}$ .

We consider the square matrix of order  $m$

$$L = \begin{pmatrix} 0 & 0 & \dots & 0 & -\alpha_m \\ 1 & 0 & \dots & 0 & -\alpha_{m-1} \\ 0 & 1 & \dots & 0 & -\alpha_{m-2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -\alpha_1 \end{pmatrix}. \quad (36)$$

It is not difficult to verify that  $g(\lambda)$  is the characteristic polynomial of  $L$ :

$$|\lambda E - L| = \begin{vmatrix} \lambda & 0 & 0 & \dots & 0 & \alpha_m \\ -1 & \lambda & 0 & \dots & 0 & \alpha_{m-1} \\ 0 & -1 & \lambda & \dots & 0 & \alpha_{m-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & \alpha_1 + \lambda \end{vmatrix} = g(\lambda).$$

On the other hand, the minor of the element  $\alpha_m$  in the characteristic determinant is equal to  $\pm 1$ . Therefore  $D_{m-1}(\lambda) = 1$  and  $i_1(\lambda) = \frac{D_m(\lambda)}{D_{m-1}(\lambda)} = D_m(\lambda) = g(\lambda)$ ,  $i_2(\lambda) = \dots = i_m(\lambda) = 1$ .

Thus,  $L$  has a single invariant polynomial different from 1, namely  $g(\lambda)$ .

We shall call  $L$  the *companion matrix* of the polynomial  $g(\lambda)$ .

Let  $A = \|a_{ik}\|_1^n$  be a matrix with the invariant polynomials

$$i_1(\lambda), i_2(\lambda), \dots, i_t(\lambda), i_{t+1}(\lambda) = 1, \dots, i_n(\lambda) = 1. \quad (37)$$

Here the polynomials  $i_1(\lambda), i_2(\lambda), \dots, i_t(\lambda)$  have positive degrees and, from the second onwards, each divides the preceding one. We denote the companion matrices of these polynomials by  $L_1, L_2, \dots, L_t$ .

Then the quasi-diagonal matrix of order  $n$

$$L_1 = \{L_1, L_2, \dots, L_t\} \quad (38)$$

has the polynomials (37) as its invariant polynomials (see Theorem 4 on p. 141). Since the matrices  $A$  and  $L_1$  have the same invariant polynomials, they are similar, i.e., there always exists a non-singular matrix  $U$  ( $|U| \neq 0$ ) such that



$$A = UL_1U^{-1} \quad (\text{I})$$

The matrix  $L_1$  is called the *first natural normal form* of the matrix  $A$ . This normal form is characterized by: 1) the quasi-diagonal form (38), 2) the special structure of the diagonal blocks (36), and 3) the additional condition: in the sequence of characteristic polynomials of the diagonal blocks every polynomial from the second onwards divides the preceding one.<sup>17</sup>

2. We now denote by

$$\chi_1(\lambda), \chi_2(\lambda), \dots, \chi_u(\lambda) \quad (39)$$

the elementary divisors of  $A = \| a_{ik} \|_1^n$  in the number field  $\mathbb{F}$ . The corresponding companion matrices will be denoted by

$$L^{(1)}, L^{(2)}, \dots, L^{(u)}.$$

Since  $\chi_j(\lambda)$  is the only elementary divisor of  $L^{(j)}$  ( $j = 1, 2, \dots, u$ ),<sup>18</sup> the quasi-diagonal matrix

$$L_{II} = \{L^{(1)}, L^{(2)}, \dots, L^{(u)}\} \quad (40)$$

has, by Theorem 5, the polynomials (39) as its elementary divisors.

The matrices  $A$  and  $L_{II}$  have the same elementary divisors in  $\mathbb{F}$ . Therefore the matrices are similar, i.e., there always exists a non-singular matrix  $V$  ( $|V| \neq 0$ ) such that

$$A = VL_{II}V^{-1}. \quad (\text{II})$$

The matrix  $L_{II}$  is called the *second natural normal form* of the matrix  $A$ . This normal form is characterized by: 1) the quasi-diagonal form (40), 2) the special structure of the diagonal blocks (36), and 3) the additional condition: the characteristic polynomial of each diagonal block is a power of an irreducible polynomial over  $\mathbb{F}$ .

*Note.* The elementary divisors of a matrix  $A$ , in contrast to the invariant polynomials, are essentially connected with the given number field  $\mathbb{F}$ . If we choose instead of the original field  $\mathbb{F}$  another number field (which also contains the elements of the given matrix  $A$ ), then the elementary divisors may change. Together with the elementary divisors, the second natural normal form of a matrix also changes.

<sup>17</sup> From the conditions 1), 2), 3) it follows automatically that the characteristic polynomials of the diagonal blocks in  $L_1$  are the invariant polynomials of the matrix  $L_1$  and, hence, of  $A$ .

<sup>18</sup>  $\chi_j(\lambda)$  is the only invariant polynomial of  $L^{(j)}$  and is at the same time a power of a polynomial irreducible over  $\mathbb{F}$ .

Suppose, for example, that  $A = \| a_{ik} \|_1^n$  is a matrix with real elements. The characteristic polynomial of the matrix then has real coefficients. But this polynomial may have complex roots. If  $\mathbb{F}$  is the field of real numbers, then among the elementary divisors there may also be powers of irreducible quadratic trinomials with real coefficients. If  $\mathbb{F}$  is the field of complex numbers, then every elementary divisor has the form  $(\lambda - \lambda_0)^p$ .

3. Let us assume now that the number field  $\mathbb{F}$  contains not only the elements of  $A$ , but also the characteristic values of the matrix.<sup>19</sup> Then the elementary divisors of  $A$  have the form<sup>20</sup>

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots, (\lambda - \lambda_u)^{p_u} \quad (p_1 + p_2 + \dots + p_u = n). \quad (41)$$

We consider one of these elementary divisors:

$$(\lambda - \lambda_0)^p$$

and associate with it the following matrix of order  $p$ :

$$\left\| \begin{array}{cccccc} \lambda_0 & 1 & 0 & \dots & 0 \\ 0 & \lambda_0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda_0 \end{array} \right\| = \lambda_0 E^{(p)} + H^{(p)}. \quad (42)$$

It is easy to verify that this matrix has only the one elementary divisor  $(\lambda - \lambda_0)^p$ . The matrix (42) will be called the *Jordan block* corresponding to the elementary divisor  $(\lambda - \lambda_0)^p$ .

The Jordan blocks corresponding to the elementary divisors (41) will be denoted by

$$J_1, J_2, \dots, J_u.$$

Then the quasi-diagonal matrix

$$J = \{ J_1, J_2, \dots, J_u \}$$

has the powers (41) as its elementary divisors.

The matrix  $J$  can also be written in the form

$$J = \{ \lambda_1 E_1 + H_1, \lambda_2 E_2 + H_2, \dots, \lambda_u E_u + H_u \};$$

where

$$E_k = E^{(p_k)}, H_k = H^{(p_k)} \quad (k = 1, 2, \dots, u).$$

<sup>19</sup> This always holds for an arbitrary matrix  $A$  if  $\mathbb{F}$  is the field of complex numbers.

<sup>20</sup> Among the numbers  $\lambda_1, \lambda_2, \dots, \lambda_u$  there may be some that are equal.

Since the matrices  $A$  and  $J$  have the same elementary divisors, they are similar, i.e., there exists a non-singular matrix  $T$  ( $|T| \neq 0$ ) such that

$$A = TJT^{-1} = T\{\lambda_1 E_1 + H_1, \lambda_2 E_2 + H_2, \dots, \lambda_u E_u + H_u\}T^{-1} \quad (\text{III})$$

The matrix  $J$  is called the *Jordan normal form* or simply *Jordan form* of  $A$ . The Jordan normal form is characterized by its quasi-diagonal form and by the special structure (42) of the diagonal blocks.

The following scheme describes the Jordan matrix  $J$  for the elementary divisors  $(\lambda - \lambda_1)^2, (\lambda - \lambda_2)^3, \lambda - \lambda_3, (\lambda - \lambda_4)^2$ :

$$J = \left\| \begin{array}{cccccccc} \lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_4 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_4 \end{array} \right\| \quad (43)$$

If (and only if) all the elementary divisors of a matrix  $A$  are of the first degree, the Jordan form is a diagonal matrix, and in this case we have:

$$A = T\{\lambda_1, \lambda_2, \dots, \lambda_n\}T^{-1}. \quad (44)$$

Thus: *A matrix  $A$  has simple structure* (see Chapter III, § 8) *if and only if all its elementary divisors are of the first degree.*<sup>21</sup>

Instead of the Jordan block (42) sometimes the 'lower' Jordan block of order  $p$  is used:

$$\left\| \begin{array}{cccccc} \lambda_0 & 0 & . & . & 0 & 0 \\ 1 & \lambda_0 & . & . & 0 & 0 \\ 0 & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & \lambda_0 & 0 \\ 0 & . & . & . & 0 & 1 & \lambda_0 \end{array} \right\| = \lambda_0 E^{(p)} + F^{(p)}$$

This matrix also has the single elementary divisor  $(\lambda - \lambda_0)^p$  only. To the elementary divisors (41) there corresponds the *lower Jordan matrix*.<sup>22</sup>

<sup>21</sup> The elementary divisors of degree 1 are often called 'linear' or 'simple' elementary divisors.

<sup>22</sup> The matrix  $J$  is often called the *upper Jordan matrix*, in contrast to the lower Jordan matrix  $J_{(1)}$ .

$$J_{(1)} = \{ \lambda_1 E_1 + F_1, \lambda_2 E_2 + F_2, \dots, \lambda_u E_u + F_u \}$$

$$(E_k = E^{(p_k)}, F_k = F^{(p_k)}; k = 1, 2, \dots, u).$$

An arbitrary matrix  $A$  having the elementary divisors (41) is always similar to  $J_{(1)}$ , i.e., there exists a non-singular matrix  $T_1$  ( $|T_1| \neq 0$ ) such that

$$A = T_1 J_{(1)} T_1^{-1} = T_1 \{ \lambda_1 E_1 + F_1, \lambda_2 E_2 + F_2, \dots, \lambda_u E_u + F_u \} T_1^{-1} \quad (\text{IV})$$

We also note that if  $\lambda_0 \neq 0$ , each of the two matrices

$$\lambda_0 (E^{(p)} + H^{(p)}), \quad \lambda_0 (E^{(p)} + F^{(p)})$$

has only the single elementary divisor  $(\lambda - \lambda_0)^p$ . Therefore for a *non-singular* matrix  $A$  having the elementary divisors (41) we have, apart from (III) and (IV), the representations

$$A = T_2 \{ \lambda_1 (E_1 + H_1), \lambda_2 (E_2 + H_2), \dots, \lambda_u (E_u + H_u) \} T_2^{-1}, \quad (\text{V})$$

$$A = T_3 \{ \lambda_1 (E_1 + F_1), \lambda_2 (E_2 + F_2), \dots, \lambda_u (E_u + F_u) \} T_3^{-1}; \quad (\text{VI})$$

## § 7. The Elementary Divisors of the Matrix $f(A)$

1. In this section we consider the following problem:

*Given the elementary divisors (in the field of complex numbers) of a matrix  $A = \| a_{ik} \|_1^n$  and given a function  $f(\lambda)$  defined on the spectrum of  $A$ , to determine the elementary divisors (in the field of complex numbers) of the matrix  $f(A)$ .*

The matrix  $f(A)$  does not alter if we replace the function  $f(\lambda)$  by a polynomial that assumes on the spectrum of  $A$  the same values as  $f(\lambda)$  (see Chapter V, § 1). Without loss of generality we may therefore assume in what follows that  $f(\lambda)$  is a polynomial.

We denote by

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots, (\lambda - \lambda_u)^{p_u}$$

the elementary divisors of  $A$ .<sup>23</sup> Thus  $A$  is similar to the Jordan matrix

$$A = T J T^{-1},$$

and so

$$f(A) = T f(J) T^{-1}$$

<sup>23</sup> Among the numbers  $\lambda_1, \lambda_2, \dots, \lambda_u$  there may be some that are equal.

Moreover,

$$J = \{J_1, J_2, \dots, J_u\}, \quad J_i = \lambda_i E^{(p_i)} + H^{(p_i)} \quad (i = 1, 2, \dots, u)$$

and

$$f(J) = \{f(J_1), f(J_2), \dots, f(J_u)\}, \quad (45)$$

where (see Example 2 on p. 100)

$$f(J_i) = \begin{vmatrix} f(\lambda_i) & \frac{f'(\lambda_i)}{1!} & \dots & \frac{f^{(p_i-1)}(\lambda_i)}{(p_i-1)!} \\ 0 & f(\lambda_i) & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \frac{f'(\lambda_i)}{1!} \\ 0 & 0 & \dots & f(\lambda_i) \end{vmatrix}. \quad (46)$$

Since the similar matrices  $f(A)$  and  $f(J)$  have the same elementary divisors, we shall from now on consider  $f(J)$  instead of  $f(A)$ .

2. Let us determine the defect<sup>24</sup>  $d$  of  $f(A)$  or, what is the same, of  $f(J)$ . The defect of a quasi-diagonal matrix is equal to the sum of the defects of the various diagonal blocks and the defect of  $f(J_i)$  (see 46)) is equal to the smaller of the numbers  $k_i$  and  $p_i$ , where  $k_i$  is the multiplicity of  $\lambda_i$  as a root of  $f(\lambda)$ ,<sup>25</sup> so that

$$f(\lambda_i) = f'(\lambda_i) = \dots = f^{(k_i-1)}(\lambda_i) = 0, \quad f^{(k_i)}(\lambda_i) \neq 0 \quad (i = 1, 2, \dots, u).$$

We have thus arrived at the following theorem:

**THEOREM 8:** *The defect of the matrix  $f(A)$ , where  $A$  has the elementary divisors*

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots, (\lambda - \lambda_u)^{p_u} \quad (47)$$

*is given by the formula*

$$d = \sum_{i=1}^u \min(k_i, p_i); \quad (48)$$

<sup>24</sup>  $d = n - r$ , where  $r$  is the rank of  $f(A)$ . If the elementary divisors of a matrix are known, then the defect of the matrix is determined as the number of elementary divisors corresponding to the characteristic value 0, i.e., as the number of elementary divisors of the form  $\lambda^w$ .

<sup>25</sup>  $k_i$  may be equal to zero; in that case  $f(\lambda_i) \neq 0$ .

here  $k_i$  is the multiplicity of  $\lambda_i$  as root of  $f(\lambda)$  ( $i=1, 2, \dots, u$ ).<sup>26</sup>

As an application of this theorem we shall determine all the elementary divisors of an arbitrary matrix  $A = \| a_{ik} \|_1^n$  that corresponds to a characteristic value  $\lambda_0$ :

$$\underbrace{\lambda - \lambda_0, \dots, \lambda - \lambda_0}_{g_1}; \underbrace{(\lambda - \lambda_0)^2, \dots, (\lambda - \lambda_0)^2}_{g_2}; \dots; \underbrace{(\lambda - \lambda_0)^m, \dots, (\lambda - \lambda_0)^m}_{g_m};$$

where  $g_i \geq 0$  ( $i=1, 2, \dots, m-1$ ),  $g_m > 0$ , provided the defects

of the matrices  $d_1, d_2, \dots, d_m$

are given,  $A - \lambda_0 E, (A - \lambda_0 E)^2, \dots, (A - \lambda_0 E)^m$

For this purpose we note that  $(A - \lambda_0 E)^j = f_j(A)$ , where  $f_j(\lambda) = (\lambda - \lambda_0)^j$  ( $j = 1, 2, \dots, m$ ). In order to determine the defect of  $(A - \lambda_0 E)^j$  we have, therefore, to set  $k_i = j$  in (48) for the elementary divisors corresponding to the characteristic value  $\lambda_0$  and  $k_i = 0$  for all the other terms ( $j = 1, 2, \dots, m$ ). Thus we obtain the formulas

$$\left. \begin{aligned} g_1 + g_2 + g_3 + \dots + g_m &= d_1, \\ g_1 + 2g_2 + 2g_3 + \dots + 2g_m &= d_2, \\ g_1 + 2g_2 + 3g_3 + \dots + 3g_m &= d_3, \\ &\vdots \\ g_1 + 2g_2 + 3g_3 + \dots + mg_m &= d_m. \end{aligned} \right\} \quad (49)$$

Hence<sup>27</sup>

$$g_j = 2d_j - d_{j-1} - d_{j+1} \quad (j = 1, 2, \dots, m; d_0 = 0, d_{m+1} = d_m). \quad (50)$$

**3.** Let us return to the basic problem of determining the elementary divisors of the matrix  $f(A)$ . As we have mentioned above, the elementary divisors of  $f(A)$  coincide with those of  $f(J)$  and the elementary divisors of a quasi-diagonal matrix coincide with those of the diagonal blocks (see Theorem 5). Therefore the problem reduces to finding the elementary divisors of a matrix  $C$  of regular triangular form:

<sup>26</sup> In the general case, where  $f(\lambda)$  is not a polynomial, then  $\min(k_i, p_i)$  in (48) has to be interpreted as the number  $p_i$  if

$$f(\lambda_i) = f'(\lambda_i) = \dots = f^{(p_i)}(\lambda_i) = 0$$

and as the number  $k_i \leq p_i$  if

$$f(\lambda_i) = f'(\lambda_i) = \dots = f^{(k_i-1)}(\lambda_i) = 0, \\ f^{(k_i)}(\lambda_i) \neq 0$$

 $(i = 1, 2, \dots, u).$ 

<sup>27</sup> The number  $m$  is characterized by the fact that  $d_{m-1} < d_m = d_{m+j}$  ( $j = 1, 2, \dots$ ).

$$C = \sum_{k=0}^{p-1} a_k H^k = \begin{vmatrix} a_0 & a_1 & \cdot & \cdot & \cdot & a_{p-1} \\ 0 & a_0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & a_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & a_0 \end{vmatrix}. \quad (51)$$

We consider separately two cases:

1.  $a_1 \neq 0$ . The characteristic polynomial of  $C$  is obviously equal to

$$D_p(\lambda) = (\lambda - a_0)^p.$$

Since  $D_{p-1}(\lambda)$  divides  $D_p(\lambda)$  without remainder, we have

$$D_{p-1}(\lambda) = (\lambda - a_0)^g \quad (g \leq p).$$

Here  $D_{p-1}(\lambda)$  denotes the greatest common divisor of the minors of order  $p-1$  in the characteristic matrix

$$\lambda E - C = \begin{vmatrix} \lambda - a_0 & -a_1 & \cdot & \cdot & \cdot & -a_{p-1} \\ 0 & \lambda - a_0 & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & -a_1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & \lambda - a_0 \end{vmatrix}.$$

It is easy to see that when the minor of the zero element marked by '+' is expanded, every term contains at least one factor  $\lambda - a_0$ , except the product of the elements on the main diagonal, which is  $(-a_1)^{p-1}$  and is therefore in our case different from zero. But since  $D_{p-1}(\lambda)$  must be a power of  $\lambda - a_0$ , we see that  $g = 0$ . But then it follows from

$$D_p(\lambda) = (\lambda - a_0)^p, \quad D_{p-1}(\lambda) = 1$$

that  $C$  has only the one elementary divisor  $(\lambda - a_0)^p$ .

2.  $a_1 = \dots = a_{k-1} = 0, a_k \neq 0$ . In this case,

$$C = a_0 E + a_k H^k + \dots + a_{p-1} H^{p-1}$$

Therefore for the positive integer  $j$  the defect of the matrix

$$(C - a_0 E)^j = a_k^j H^{kj} + \dots$$

is given by

$$d_j = \begin{cases} kj, & \text{when } kj \leq p, \\ p, & \text{when } kj > p. \end{cases}$$

We set

$$p = qk + h \quad (0 \leq h < k). \quad (52)$$

Then<sup>28</sup>

$$d_1 = k, d_2 = 2k, \dots, d_q = qk, d_{q+1} = p. \quad (53)$$

Therefore we have by (50)

$$g_1 = \dots = g_{q-1} = 0, \quad g_q = k - h, \quad g_{q+1} = h.$$

Thus, the matrix  $C$  has the elementary divisors

$$\underbrace{(\lambda - a_0)^{q+1}, \dots, (\lambda - a_0)^{q+1}}_h, \quad \underbrace{(\lambda - a_0)^q, \dots, (\lambda - a_0)^q}_{k-h}, \quad (54)$$

where the integers  $q > 0$  and  $h \geq 0$  are determined by (52).

4. Now we are in a position to ascertain what elementary divisors the matrix  $f(J)$  has (see (45) and (46)). To each elementary divisor of  $A$

$$(\lambda - \lambda_0)^p$$

there corresponds in  $f(J)$  the diagonal cell

$$f(\lambda_0 E + H) = \sum_{i=0}^{p-1} \frac{f^{(i)}(\lambda_0)}{i!} H^i = \begin{vmatrix} f(\lambda_0) & \frac{f'(\lambda_0)}{1!} & \dots & \frac{f^{(p-1)}(\lambda_0)}{(p-1)!} \\ 0 & f(\lambda_0) & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \frac{f'(\lambda_0)}{1!} \\ 0 & 0 & \dots & f(\lambda_0) \end{vmatrix}. \quad (55)$$

Clearly the problem reduces to finding the elementary divisors of a cell of the form (55). But the matrix (55) is of the regular triangular form (51), where

$$a_0 = f(\lambda_0), \quad a_1 = f'(\lambda_0), \quad a_2 = \frac{f''(\lambda_0)}{2!}, \quad \dots$$

Thus we arrive at the theorem:

<sup>28</sup> In this case the number  $q + 1$  plays the role of  $m$  in (49) and (50).



THEOREM 9: *The elementary divisors of the matrix  $f(A)$  are obtained from those of  $A$  in the following way: To an elementary divisor*

$$(\lambda - \lambda_0)^p \quad (56)$$

*of  $A$  for  $p=1$  or for  $p > 1$  and  $f'(\lambda_0) \neq 0$  there corresponds a single elementary divisor*

$$(\lambda - f(\lambda_0))^p \quad (57)$$

*of  $f(A)$ ; for  $p > 1$  and  $f(\lambda_0) = \dots = f^{(k-1)}(\lambda_0) = 0$ ,  $f^{(k)}(\lambda_0) \neq 0$  ( $k < p$ ) to the elementary divisor (56) of  $A$  there correspond the following elementary divisors of  $f(A)$ :*

$$\underbrace{(\lambda - f(\lambda_0))^{q+1}, \dots, (\lambda - f(\lambda_0))^{q+1}}_h, \quad \underbrace{(\lambda - f(\lambda_0))^q, \dots, (\lambda - f(\lambda_0))^q}_{k-h} \quad (58)$$

where

$$p = qk + h, \quad 0 \leq q, \quad 0 \leq h < k;$$

finally, for  $p > 1$ ,  $f'(\lambda_0) = \dots = f^{(p-1)}(\lambda_0) = 0$ , to the elementary divisor (56) there correspond  $p$  elementary divisors of the first degree of  $f(A)$ :<sup>29</sup>

$$\lambda - f(\lambda_0), \dots, \lambda - f(\lambda_0). \quad (59)$$

We note the following special cases of this theorem.

1. If  $\lambda_1, \lambda_2, \dots, \lambda_n$  are the characteristic values of  $A$ , then  $f(\lambda_1), f(\lambda_2), \dots, f(\lambda_n)$  are the characteristic values of  $f(A)$ . (In both sequences each characteristic value is repeated as often as its multiplicity as a root of the characteristic equation indicates.)<sup>30</sup>

2. If the derivative  $f'(\lambda)$  is not zero on the spectrum of  $A$ ,<sup>31</sup> then in going from  $A$  to  $f(A)$  the elementary divisors are not 'split up' i.e., if  $A$  has the elementary divisors

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots, (\lambda - \lambda_n)^{p_n},$$

then  $f(A)$  has the elementary divisors

$$(\lambda - f(\lambda_1))^{p_1}, (\lambda - f(\lambda_2))^{p_2}, \dots, (\lambda - f(\lambda_n))^{p_n}$$

<sup>29</sup> (57) is obtained from (58) by setting  $k=1$ ; (59) is obtained from (58) by setting  $k=p$  or  $k > p$ .

<sup>30</sup> Statement 1. was established separately in Chapter IV, p. 84.

<sup>31</sup> I.e.,  $f'(\lambda_i) \neq 0$  for those  $\lambda_i$  that are multiple roots of the minimal polynomial.

## § 8. A General Method of Constructing the Transforming Matrix

In many problems in the theory of matrices and its applications it is sufficient to know the normal form into which a given matrix  $A = \| a_{ik} \|_1^n$  can be carried by similarity transformations. The normal form is completely determined by the invariant polynomials of the characteristic matrix  $\lambda E - A$ . To find the latter, we can use the defining formulas (see (10) on p. 139) or the reduction of the characteristic matrix  $\lambda E - A$  to canonical diagonal form by elementary transformations.

In some problems, however, it is necessary to know not only the normal form  $\tilde{A}$  of the given matrix  $A$ , but also a non-singular transforming matrix  $T$ .

1. An immediate method of determining  $T$  consists in the following. The equation

$$A = T\tilde{A}T^{-1}$$

can be written as:

$$AT - T\tilde{A} = 0.$$

This matrix equation in  $T$  is equivalent to a system of  $n^2$  linear homogeneous equations in the  $n^2$  unknown coefficients of  $T$ . The determination of a transforming matrix reduces to the solution of this system of  $n^2$  equations. Moreover, we have to choose from the set of all solutions one for which  $|T| \neq 0$ . The existence of such a solution is certain, since  $A$  and  $\tilde{A}$  have the same invariant polynomials.<sup>32</sup>

Note that whereas the normal form is uniquely determined by the matrix  $A$ ,<sup>33</sup> for the transforming matrix  $T$  we always have an innumerable set of values that are given by

$$T = UT_1, \quad (60)$$

where  $T_1$  is one of the transforming matrices and  $U$  is an arbitrary matrix that is permutable with  $A$ .<sup>34</sup>

<sup>32</sup> From this fact follows the similarity of  $\tilde{A}$  and  $A$ .

<sup>33</sup> This statement is unconditionally true as regards the first natural normal form. As far as the second normal form or the Jordan normal form is concerned, they are uniquely determined to within the order of the diagonal blocks.

<sup>34</sup> The formula (60) may be replaced by

$$T = T_1V,$$

where  $V$  is an arbitrary matrix permutable with  $\tilde{A}$ .

The method proposed above for determining a transforming matrix  $T$  is simple enough in concept but of little use in practice, since it requires a great many computations (even for  $n=4$  we have to solve 16 linear equations).

**2.** We proceed to explain a more efficient method of constructing the transforming matrix  $T$ . This method is based on the Supplement to Theorem 7 (p. 148). According to this, we can choose as the transforming matrix

$$T = Q(\tilde{A}), \quad (61)$$

provided

$$\lambda E - \tilde{A} = P(\lambda)(\lambda E - A)Q(\lambda).$$

The latter equation expresses the equivalence of the characteristic matrices  $\lambda E - A$  and  $\lambda E - \tilde{A}$ . Here  $P(\lambda)$  and  $Q(\lambda)$  are polynomial matrices with constant non-zero determinants.

For the actual process of finding  $Q(\lambda)$  we reduce the two  $\lambda$ -matrices  $\lambda E - A$  and  $\lambda E - \tilde{A}$  to canonical form by means of the corresponding elementary transformations

$$\{i_n(\lambda), i_{n-1}(\lambda), \dots, i_1(\lambda)\} = P_1(\lambda)(\lambda E - A)Q_1(\lambda) \quad (62)$$

$$\{i_n(\lambda), i_{n-1}(\lambda), \dots, i_1(\lambda)\} = P_2(\lambda)(\lambda E - \tilde{A})Q_2(\lambda) \quad (63)$$

where

$$Q_1(\lambda) = T_1 T_2 \dots T_p, \quad Q_2(\lambda) = T_1^* T_2^* \dots T_p^*, \quad (64)$$

and where  $T_1, \dots, T_p, T_1^*, \dots, T_p^*$  are the elementary matrices corresponding to the elementary operations on the columns of the  $\lambda$ -matrices  $\lambda E - A$  and  $\lambda E - \tilde{A}$ . From (62), (63), and (64) it follows that

$$\lambda E - \tilde{A} = P(\lambda)(\lambda E - A)Q(\lambda),$$

where

$$Q(\lambda) = Q_1(\lambda)Q_2^{-1}(\lambda) = T_1 T_2 \dots T_p T_p^{*-1} T_{p-1}^{*-1} \dots T_1^{*-1}. \quad (65)$$

We can compute the matrix  $Q(\lambda)$  by applying successively to the columns of the unit matrix  $E$  the elementary operations with the matrices  $T_1, \dots, T_p, T_p^{*-1}, \dots, T_1^{*-1}$ . After this (in accordance with (61)) we replace the argument  $\lambda$  in  $Q(\lambda)$  by the matrix  $\tilde{A}$ .

*Example.*

$$A = \begin{vmatrix} 1 & -1 & 1 & -1 \\ -3 & 3 & -5 & 4 \\ 8 & -4 & 3 & -4 \\ 15 & -10 & 11 & -11 \end{vmatrix}.$$

Let us introduce a symbolic notation for the right elementary operations and the corresponding matrices (see pp. 130-131):

$$T'' = [(c) i], T''' = [i + (b(\lambda)) j], T'''' = [ij].$$

In transforming the characteristic matrix  $\lambda E - A$  into normal diagonal form we shall at the same time keep a record of the elementary right operations to be performed, i.e., the operations on the columns:

$$\begin{aligned} \lambda E - A = & \left\| \begin{array}{cccc} \lambda - 1 & 1 & -1 & 1 \\ 3 & \lambda - 3 & 5 & -4 \\ -8 & 4 & \lambda - 3 & 4 \\ -15 & 10 & -11 & \lambda + 11 \end{array} \right\|, \left\| \begin{array}{cccc} 0 & 0 & 0 & 1 \\ 4\lambda - 1 & \lambda + 1 & 1 & -4 \\ -4\lambda - 4 & 0 & \lambda + 1 & 4 \\ -\lambda^2 - 10\lambda - 4 & -\lambda - 1 & \lambda & \lambda + 11 \end{array} \right\|, \\ & \left\| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ -4 & \lambda + 1 & 1 & 4\lambda - 1 \\ 4 & 0 & \lambda + 1 & -4\lambda - 4 \\ \lambda + 11 & -\lambda - 1 & \lambda & -\lambda^2 - 10\lambda - 4 \end{array} \right\|; \\ & \left\| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & \lambda + 1 & -1 & 4\lambda - 1 \\ 0 & 0 & \lambda + 1 & -4\lambda - 4 \\ 0 & -\lambda - 1 & \lambda & -\lambda^2 - 10\lambda - 4 \end{array} \right\|, \left\| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -\lambda^2 - 2\lambda - 1 & \lambda + 1 & -4\lambda^2 - 7\lambda - 3 \\ 0 & -\lambda^2 - 2\lambda - 1 & \lambda & -\lambda^2 - 9\lambda - 4 \end{array} \right\|, \\ & \left\| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & \lambda + 1 & -\lambda^2 - 2\lambda - 1 & -4\lambda^2 - 7\lambda - 3 \\ 0 & \lambda & -\lambda^2 - 2\lambda - 1 & -5\lambda^2 - 9\lambda - 4 \end{array} \right\|; \\ & \left\| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \lambda^2 + 2\lambda + 1 & 4\lambda^2 + 7\lambda + 3 \\ 0 & 0 & \lambda^2 + 2\lambda + 1 & 5\lambda^2 + 9\lambda + 4 \end{array} \right\|, \left\| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \lambda^2 + 2\lambda + 1 & -\lambda^2 - 3\lambda - 2 \\ 0 & 0 & \lambda^2 + 2\lambda + 1 & -\lambda - 1 \end{array} \right\|, \\ & \left\| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -\lambda^2 - 2\lambda - 1 & \lambda + 1 \\ 0 & 0 & \lambda^2 + 2\lambda + 1 & -\lambda^2 - 3\lambda - 2 \end{array} \right\|; \\ & \left\| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \lambda + 1 & -\lambda^2 - 2\lambda - 1 \\ 0 & 0 & -\lambda^2 - 3\lambda - 2 & \lambda^2 + 2\lambda + 1 \end{array} \right\|, \left\| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \lambda + 1 & 0 \\ 0 & 0 & -\lambda^2 - 3\lambda - 2 & -(\lambda + 1)^2 \end{array} \right\|, \\ & \left\| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \lambda + 1 & 0 \\ 0 & 0 & 0 & (\lambda + 1)^2 \end{array} \right\|. \end{aligned}$$

Here

$$Q_1(\lambda) = [1 + (1 - \lambda) 4] [2 - 4] [3 + 4] [14] [2 - (\lambda + 1) 3] [4 + (1 - 4\lambda) 3] [23] \times \\ \times [4 - (5) 3] [43] [4 + (\lambda + 1) 3].$$

We have found the invariant polynomials  $(\lambda + 1)^3$ ,  $(\lambda + 1)$ , 1, and 1 of  $A$ . The matrix has two elementary divisors,  $(\lambda + 1)^3$  and  $(\lambda + 1)$ . Therefore the Jordan normal form is

$$J = \begin{vmatrix} - & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix}.$$

By elementary operations we bring the matrix  $\lambda E - J$  into normal diagonal form

$$\begin{aligned} \lambda E - J = & \begin{vmatrix} \lambda + 1 & -1 & 0 & 0 \\ 0 & \lambda + 1 & -1 & 0 \\ 0 & 0 & \lambda + 1 & 0 \\ 0 & 0 & 0 & \lambda + 1 \end{vmatrix}, \begin{vmatrix} \lambda + 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & (\lambda + 1)^2 & \lambda + 1 & 0 \\ 0 & 0 & 0 & \lambda + 1 \end{vmatrix}, \\ & \begin{vmatrix} 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ (\lambda + 1)^3 & (\lambda + 1)^2 & \lambda + 1 & 0 \\ 0 & 0 & 0 & \lambda + 1 \end{vmatrix}; \\ & \begin{vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ (\lambda + 1)^3 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda + 1 \end{vmatrix}, \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & (\lambda + 1)^3 \\ 0 & 0 & \lambda + 1 & 0 \end{vmatrix}, \\ & \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \lambda + 1 & 0 \\ 0 & 0 & 0 & (\lambda + 1)^3 \end{vmatrix}. \end{aligned}$$

Here

$$Q_2(\lambda) = [2 + (\lambda + 1) 3] [1 + (\lambda + 1) 2] [12] [23] [34].$$

Therefore

$$\begin{aligned} Q(\lambda) &= Q_1(\lambda) Q_2^{-1}(\lambda) \\ &= [1 + (1 - \lambda) 4] [2 - 4] [3 + 4] [14] [2 - (\lambda + 1) 3] [4 + (1 - \lambda) 3] [23] [4 - (5) 3] \\ &\quad \times [43] [4 + (\lambda + 1) 3] [34] [23] [12] [1 - (\lambda + 1) 2] [2 - (\lambda + 1) 3]. \end{aligned}$$

We apply these elementary operations successively to the unit matrix  $E$ :

$$\begin{aligned} E = & \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}, \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 - \lambda & -1 & 1 & 1 \end{vmatrix}, \\ & \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 - 1 & 1 & 1 - \lambda & \end{vmatrix}, \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & -\lambda - 1 & 1 & 0 \\ 1 & -\lambda - 2 & 1 & 1 - \lambda \end{vmatrix}; \end{aligned}$$

$$\begin{aligned}
 & \left\| \begin{array}{cccc} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & -\lambda-1 & 1 & 1-4\lambda \\ 1 & -\lambda-2 & 1 & 2-5\lambda \end{array} \right\|, \left\| \begin{array}{cccc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & -\lambda-1 & 1-4\lambda \\ 1 & 1 & -\lambda-2 & 2-5\lambda \end{array} \right\|, \left\| \begin{array}{cccc} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -5 \\ 0 & 1 & -\lambda-1 & \lambda+6 \\ 1 & 1 & -\lambda-2 & 12 \end{array} \right\|, \\
 & \left\| \begin{array}{cccc} 0 & 0 & 1 & 0 \\ 0 & 0 & -5 & 1 \\ 0 & 1 & \lambda+6 & -\lambda-1 \\ 1 & 1 & 12 & -\lambda-2 \end{array} \right\|, \left\| \begin{array}{cccc} 0 & 0 & 1 & \lambda+1 \\ 0 & 0 & -5 & -5\lambda-4 \\ 0 & 1 & \lambda+6 & \lambda^2+6\lambda+5 \\ 1 & 1 & 12 & 11\lambda+10 \end{array} \right\|, \left\| \begin{array}{cccc} \lambda+1 & 0 & 0 & 1 \\ -5\lambda-4 & 0 & 0 & -5 \\ \lambda^2+6\lambda+5 & 0 & 1 & \lambda+6 \\ 11\lambda+10 & 1 & 1 & 12 \end{array} \right\|, \\
 & \left\| \begin{array}{cccc} \lambda+1 & 0 & 0 & 1 \\ -5\lambda-4 & 0 & 0 & -5 \\ \lambda^2+6\lambda+5 & -\lambda-1 & 1 & \lambda+6 \\ 10\lambda+9 & -\lambda & 1 & 12 \end{array} \right\|.
 \end{aligned}$$

Thus

$$\begin{aligned}
 Q(\lambda) &= \left\| \begin{array}{cccc} \lambda+1 & 0 & 0 & 1 \\ -5\lambda-4 & 0 & 0 & -5 \\ \lambda^2+6\lambda+5 & -\lambda-1 & 1 & \lambda+6 \\ 10\lambda+9 & -\lambda & 1 & 12 \end{array} \right\| \\
 &= \left\| \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right\| \lambda^2 + \left\| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ -5 & 0 & 0 & 0 \\ 6 & -1 & 0 & 1 \\ 10 & -1 & 0 & 0 \end{array} \right\| \lambda + \left\| \begin{array}{cccc} 1 & 0 & 0 & 1 \\ -4 & 0 & 0 & -5 \\ 5 & -1 & 1 & 6 \\ 9 & 0 & 1 & 12 \end{array} \right\|.
 \end{aligned}$$

Observing that

$$J^2 = \left\| \begin{array}{cccc} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right\|,$$

we have

$$\begin{aligned}
 T = Q(J) &= \left\| \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right\| \cdot \left\| \begin{array}{cccc} 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right\| + \\
 &+ \left\| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ -5 & 0 & 0 & 0 \\ 6 & -1 & 0 & 1 \\ 12 & -1 & 0 & 0 \end{array} \right\| \cdot \left\| \begin{array}{cccc} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{array} \right\| + \left\| \begin{array}{cccc} 1 & 0 & 0 & 1 \\ -4 & 0 & 0 & -5 \\ 5 & -1 & 1 & 6 \\ 9 & 0 & 1 & 12 \end{array} \right\| \\
 &= \left\| \begin{array}{cccc} 0 & 1 & 0 & 1 \\ 1 & -5 & 0 & -5 \\ 0 & 4 & 1 & 5 \\ -1 & 11 & 0 & 12 \end{array} \right\|.
 \end{aligned}$$

Check:

$$AT = \begin{vmatrix} 0 & -1 & 1 & -1 \\ -1 & 6 & -5 & 5 \\ 0 & -4 & 3 & -5 \\ 1 & -12 & 11 & -12 \end{vmatrix}, \quad TJ = \begin{vmatrix} 0 & -1 & 1 & -1 \\ -1 & 6 & -5 & 5 \\ 0 & -4 & 3 & -5 \\ 1 & -12 & 11 & -12 \end{vmatrix},$$

i.e.,  $AT = TJ$ .

$$|T| = \begin{vmatrix} 0 & 1 & 0 & 1 \\ 1 & -5 & 0 & -5 \\ 0 & 4 & 1 & 5 \\ -1 & 11 & 0 & 12 \end{vmatrix} = -1 \neq 0.$$

Therefore

$$A = TJT^{-1}.$$

### § 9. Another Method of Constructing a Transforming Matrix

1. We shall now explain another method of constructing a transforming matrix which often leads to fewer computations than the method of the preceding section. However, we shall apply this second method only when the Jordan normal form and the elementary divisors

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots \quad (66)$$

of the given matrix  $A$  are known.

Let  $A = TJT^{-1}$ , where

$$J = \{\lambda_1 E^{(p_1)} + H^{(p_1)}, \lambda_2 E^{(p_2)} + H^{(p_2)}, \dots\} = \begin{vmatrix} \overbrace{\lambda_1 & 1 & \dots & 0}^{p_1} & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \ddots & 1 \\ 0 & \dots & & \lambda_1 & & \\ & & & & \overbrace{\lambda_2 & 1 & \dots & 0}^{p_2} & \\ & & & & \ddots & \ddots & \ddots & \\ & & & & & \ddots & \ddots & 1 \\ & & & & 0 & \dots & & \lambda_2 \end{vmatrix}.$$

Then denoting the  $k$ -th column of  $T$  by  $t_k$  ( $k = 1, 2, \dots, n$ ), we replace the matrix equation

$$AT = TJ$$

by the equivalent system of equations

$$At_1 = \lambda_1 t_1, At_2 = \lambda_1 t_2 + t_1, \dots, At_{p_1} = \lambda_1 t_{p_1} + t_{p_1-1} \quad (67)$$

$$At_{p_1+1} = \lambda_2 t_{p_1+1}, At_{p_1+2} = \lambda_2 t_{p_1+2} + t_{p_1+1}, \dots, At_{p_1+p_2} = \lambda_2 t_{p_1+p_2} + t_{p_1+p_2-1} \quad (68)$$

.....

which we rewrite as follows:

$$(A - \lambda_1 E) t_1 = 0, (A - \lambda_1 E) t_2 = t_1, \dots, (A - \lambda_1 E) t_{p_1} = t_{p_1-1} \quad (67')$$

$$(A - \lambda_2 E) t_{p_1+1} = 0, (A - \lambda_2 E) t_{p_1+2} = t_{p_1+1}, \dots, (A - \lambda_2 E) t_{p_1+p_2} = t_{p_1+p_2-1} \quad (68')$$

.....

Thus, all the columns of  $T$  are split into 'Jordan chains' of columns:

$[t_1, t_2, \dots, t_{p_1}], [t_{p_1+1}, t_{p_1+2}, \dots, t_{p_1+p_2}], \dots$

To every Jordan block of  $J$  (or, what is the same, to every elementary divisor (66)) there corresponds its Jordan chain of columns. Each Jordan chain of columns is characterized by a system of equations of type (67'), (68'), etc.

The task of finding a transforming matrix  $T$  reduces to that of finding the Jordan chains that would give in all  $n$  linearly independent columns.

We shall show that these Jordan chains of columns can be determined by means of the reduced adjoint matrix  $C(\lambda)$  (see Chapter IV, § 6).

For the matrix  $C(\lambda)$  we have the identity

$$(\lambda E - A) C(\lambda) = \psi(\lambda) E. \quad (69)$$

where  $\psi(\lambda)$  is the minimal polynomial of  $A$ .

Let

$$\psi(\lambda) = (\lambda - \lambda_0)^m \chi(\lambda) \quad (\chi(\lambda_0) \neq 0).$$

We differentiate the identity (69) term by term  $m - 1$  times:

$$\left. \begin{aligned} (\lambda E - A) C'(\lambda) + C(\lambda) &= \psi'(\lambda) E \\ (\lambda E - A) C''(\lambda) + 2C'(\lambda) &= \psi''(\lambda) E \\ &\dots \dots \dots \\ (\lambda E - A) C^{(m-1)}(\lambda) + (m-1) C^{(m-2)}(\lambda) &= \psi^{(m-1)}(\lambda) E. \end{aligned} \right\} \quad (70)$$

Substituting  $\lambda_0$  for  $\lambda$  in (69) and (70) and observing that the right-hand sides are zero, we obtain

$$(A - \lambda_0 E) C = 0, (A - \lambda_0 E) D = C, (A - \lambda_0 E) F = D, \dots, (A - \lambda_0 E) K = G; \quad (71)$$

where

$$\left. \begin{aligned} C &= C(\lambda_0), D = \frac{1}{1!} C'(\lambda_0), F = \frac{1}{2!} C''(\lambda_0), \dots, G = \frac{1}{(m-2)!} C^{(m-2)}(\lambda_0) \\ K &= \frac{1}{(m-1)!} C^{(m-1)}(\lambda_0). \end{aligned} \right\} \quad (72)$$



In (71) we replace the matrices (72) by their  $k$ -th columns ( $k = 1, 2, \dots, n$ ). We obtain:

$$(A - \lambda_0 E) C_k = 0, (A - \lambda_0 E) D_k = C_k, \dots, (A - \lambda_0 E) K_k = G_k \quad (73) \\ (k = 1, 2, \dots, n).$$

Since  $C = C(\lambda_0) \neq 0$ ,<sup>35</sup> we can find a  $k$  ( $\leq n$ ) such that

$$C_k \neq 0. \quad (74)$$

Then the  $m$  columns

$$C_k, D_k, F_k, \dots, G_k, K_k \quad (75)$$

are linearly independent. For let

$$\gamma C_k + \delta D_k + \dots + \kappa K_k = 0. \quad (76)$$

Multiplying both sides of (76) successively by  $A - \lambda_0 E, \dots, (A - \lambda_0 E)^{m-1}$ , we obtain

$$\delta C_k + \dots + \kappa G_k = 0, \dots, \kappa C_k = 0. \quad (77)$$

From (76) and (77) we find by (74):

$$\gamma = \delta = \dots = \kappa = 0.$$

Since the linearly independent columns (75) satisfy the system of equations (73), they form a Jordan chain of vectors corresponding to the elementary divisor  $(\lambda - \lambda_0)^m$  (compare (73) with (67')).

If  $C_k = 0$  for some  $k$ , but  $D_k \neq 0$ , then the columns  $D_k, \dots, G_k, K_k$  form a Jordan chain of  $m - 1$  vectors, etc.

2. We shall now show first of all how to construct a transforming matrix  $T$  in the case where the elementary divisors of  $A$  are pairwise co-prime:

$$(\lambda - \lambda_1)^{m_1}, \dots, (\lambda - \lambda_s)^{m_s}, \\ (\lambda_i \neq \lambda_j \text{ for } i \neq j; i, j = 1, 2, \dots, s).$$

With the elementary divisor  $(\lambda - \lambda_i)^{m_i}$  we associate the Jordan chain of columns

$$C^{(i)}, D^{(i)}, \dots, G^{(i)}, K^{(i)},$$

constructed as indicated above. Then

$$(A - \lambda_j E) C^{(i)} = 0, (A - \lambda_j E) D^{(i)} = C^{(i)}, \dots, (A - \lambda_j E) K^{(i)} = G^{(i)}. \quad (78)$$

When we give to  $j$  the values  $1, 2, \dots, s$ , we obtain  $s$  Jordan chains containing  $n$  columns in all. These columns are linearly independent.

<sup>35</sup> From  $C(\lambda_0) = 0$  it would follow that all the elements of  $C(\lambda)$  have a common divisor of positive degree, in contradiction to the definition of  $C(\lambda)$ .

For, suppose that

$$\sum_{j=1}^s [\gamma_j C^{(j)} + \delta_j D^{(j)} + \dots + \kappa_j K^{(j)}] = 0. \quad (79)$$

We multiply both sides of (79) on the left by

$$(A - \lambda_1 E)^{m_1} \dots (A - \lambda_{j-1} E)^{m_{j-1}} (A - \lambda_j E)^{m_j-1} (A - \lambda_{j+1} E)^{m_{j+1}} \dots (A - \lambda_s E)^{m_s} \quad (80)$$

and obtain

$$\kappa_j = 0.$$

Replacing  $m_j - 1$  successively by  $m_j - 2, m_j - 3, \dots$  in (80), we find:

$$\gamma_j = \delta_j = \dots = \kappa_j = 0 \quad (j = 1, 2, \dots, s),$$

and this is what we had to prove.

We define the matrix  $T$  by the formula

$$T = (C^{(1)}, D^{(1)}, \dots, K^{(1)}; C^{(2)}, D^{(2)}, \dots, K^{(2)}; \dots; C^{(s)}, D^{(s)}, \dots, K^{(s)}). \quad (81)$$

*Example.*

$$A = \begin{pmatrix} 3 & -1 & -4 & 2 \\ 2 & 3 & -2 & -4 \\ 2 & -1 & -3 & 2 \\ 1 & 2 & -1 & -3 \end{pmatrix}; \quad \begin{aligned} \psi(\lambda) &= \Delta(\lambda) = (\lambda - 1)^2 (\lambda + 1)^2 = \lambda^4 - 2\lambda^2 + 1, \\ \text{elementary divisors: } &(\lambda - 1)^2, (\lambda + 1)^2, \\ \Psi(\lambda, \mu) &= \frac{\psi(\mu) - \psi(\lambda)}{\mu - \lambda} = \mu^3 + \lambda\mu^2 + (\lambda^2 - 2)\mu + \lambda^3 - 2\lambda, \end{aligned}$$

$$\begin{pmatrix} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 3 & 2 & 2 & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 1 & 4 & 0 & 2 & \cdot & \cdot \end{pmatrix}$$

$$C(\lambda) = \Psi(\lambda E, A) = A^3 + \lambda A^2 + (\lambda^2 - 2)A + (\lambda^3 - 2\lambda)E.$$

We make up the first column  $C_1(\lambda)$ :

$$C_1(\lambda) = [A^3]_1 + \lambda [A^2]_1 + (\lambda^2 - 2)A_1 + (\lambda^3 - 2\lambda)E_1.$$

For the computation of the first column of  $A^2$  we multiply all the rows of  $A$  into the first column of  $A$ . We obtain:<sup>36</sup>  $[A^2]_1 = (1, 4, 0, 2)$ . Multiplying all the rows of  $A$  into this column, we find:  $[A^3]_1 = (3, 6, 2, 3)$ .

Therefore

$$C_1(\lambda) = \begin{pmatrix} 3 \\ 6 \\ 2 \\ 3 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 4 \\ 0 \\ 2 \end{pmatrix} + (\lambda^2 - 2) \begin{pmatrix} 3 \\ 2 \\ 2 \\ 1 \end{pmatrix} + (\lambda^3 - 2\lambda) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \lambda^3 + 3\lambda^2 - \lambda - 3 \\ 2\lambda^2 + 4\lambda + 2 \\ 2\lambda^2 - 2 \\ \lambda^2 + 2\lambda + 1 \end{pmatrix}.$$

<sup>36</sup> The columns into which we multiply the rows are written underneath the rows of  $A$ . The elements of the row of column-sums are set up in italics, for checking.

Hence  $C_1(1) = (0, 8, 0, 4)$  and  $C'_1(1) = (8, 8, 4, 4)$ . As  $C_1(-1) = (0, 0, 0, 0)$ , we pass on to the second column and, proceeding as before, we find:  $C_2(-1) = (-4, 0, -4, 0)$  and  $C'_2(-1) = (4, -4, 4, -4)$ . We set up the matrix:

$$(C_1(1), C'_1(1); C_2(-1), C'_2(-1)) = \begin{vmatrix} 0 & 8 & -4 & 4 \\ 8 & 8 & 0 & -4 \\ 0 & 4 & -4 & 4 \\ 4 & 4 & 0 & -4 \end{vmatrix}.$$

We cancel<sup>37</sup> 4 in the first two columns and  $-4$  in the last two columns.

$$T = \begin{vmatrix} 0 & 2 & 1 & -1 \\ 2 & 2 & 0 & 1 \\ 0 & 1 & 1 & -1 \\ 1 & 1 & 0 & 1 \end{vmatrix}.$$

We leave it to the reader to verify that

$$AT = T \cdot \begin{vmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{vmatrix}$$

**3.** Coming now to the general case, we shall investigate the Jordan chains of vectors corresponding to a characteristic value  $\lambda_0$  for which there are  $p$  elementary divisors  $(\lambda - \lambda_0)^m$ ,  $q$  elementary divisors  $(\lambda - \lambda_0)^{m-1}$ ,  $r$  elementary divisors  $(\lambda - \lambda_0)^{m-2}$ , etc.

As a preliminary to this, we establish some properties of the matrices

$$C = C(\lambda_0), D = C'(\lambda_0), F = \frac{1}{2!} C''(\lambda_0), \dots, K = \frac{1}{(m-1)!} C^{(m-1)}(\lambda_0). \quad (82)$$

1. The matrices (82) can be represented in the form of polynomials in  $A$ :

$$C = h_1(A), D = h_2(A), \dots, K = h_m(A), \quad (83)$$

where

$$h_i(\lambda) = \frac{\psi(\lambda)}{(\lambda - \lambda_0)^i} \quad (i = 1, 2, \dots, m). \quad (84)$$

For

$$C(\lambda) = \psi'(\lambda E, A),$$

where

$$\psi'(\lambda, \mu) = \frac{\psi(\mu) - \psi(\lambda)}{\mu - \lambda}.$$

<sup>37</sup> A Jordan chain remains a Jordan chain when all its columns are multiplied by a number  $c \neq 0$ .

Therefore

$$\frac{1}{k!} C^{(k)}(\lambda_0) = \frac{1}{k!} \Psi^{(k)}(\lambda_0 E, A), \quad (85)$$

where

$$\begin{aligned} \frac{1}{k!} \Psi^{(k)}(\lambda_0, \mu) &= \frac{1}{k!} \left[ \frac{\partial^k}{\partial \lambda^k} \Psi(\lambda, \mu) \right]_{\lambda=\lambda_0} \\ &= \frac{1}{k!} \left[ \frac{\partial^k}{\partial \lambda^k} \frac{\Psi(\mu)}{\mu - \lambda} \right]_{\lambda=\lambda_0} = \frac{\Psi(\mu)}{(\mu - \lambda_0)^{k+1}} = h_{k+1}(\mu). \end{aligned} \quad (86)$$

(83) follows from (82), (85), and (86).

2. The matrices (82) have the ranks

$$p, 2p+q, 3p+2q+r, \dots$$

This property of the matrices (82) follows immediately from 1. and Theorem 8 (§ 7), if we equate the rank to  $n-d$  and use formula (48) for the defect of a function on  $A$  (p. 154).

3. In the sequence of matrices (82) every column of each matrix is a linear combination of the columns of every following matrix.

Let us take two matrices  $h_i(A)$  and  $h_k(A)$  in (82) (see 1.). Suppose that  $i < k$ . Then it follows from (84) that:

$$h_i(A) = h_k(A) (A - \lambda_0 E)^{k-i}.$$

Hence the  $j$ -th column  $y_j$  ( $j = 1, 2, \dots, n$ ) of  $h_i(A)$  is expressed linearly by the columns  $z_1, z_2, \dots, z_n$  of  $h_k(A)$ :

$$y_j = \sum_{\sigma=1}^n \alpha_{\sigma} z_{\sigma},$$

where  $\alpha_1, \alpha_2, \dots, \alpha_n$  are the elements of the  $j$ -th column of  $(A - \lambda_0 E)^{k-i}$ .

4. Without changing the basic formulas (71) we may replace any column in  $C$  by an arbitrary linear combination of all the columns, provided we make the corresponding replacements in  $D, \dots, K$ .

We now proceed to the construction of the Jordan chains of columns for the elementary divisors

$$\underbrace{(\lambda - \lambda_0)^m, \dots, (\lambda - \lambda_0)^1}_p; \quad \underbrace{(\lambda - \lambda_0)^{m-1}, \dots, (\lambda - \lambda_0)^1}_{q}; \dots$$

Using the properties 2. and 4., we transform the matrix  $C$  into the form

$$C = (C_1, C_2, \dots, C_p, o, o, \dots, o); \quad (87)$$

where the columns  $C_1, C_2, \dots, C_p$  are linearly independent. Now

$$D = (D_1, D_2, \dots, D_p; D_{p+1}, \dots, D_n).$$

By 3., for every  $i$  ( $1 \leq i \leq p$ )  $C_i$  is a linear combination of the columns  $D_1, D_2, \dots, D_n$ :

$$C_i = \alpha_1 D_1 + \dots + \alpha_p D_p + \alpha_{p+1} D_{p+1} + \dots + \alpha_n D_n. \quad (88)$$

We multiply both sides of this equation by  $A - \lambda_0 E$ . Observing (see (73)) that

$$(A - \lambda_0 E) C_i = 0 \quad (i=1, 2, \dots, p), \quad (A - \lambda_0 E) D_j = C_j \quad (j=1, 2, \dots, n),$$

we obtain by (87)

$$0 = \alpha_1 C_1 + \alpha_2 C_2 + \dots + \alpha_p C_p;$$

hence in (88)

$$\alpha_1 = \dots = \alpha_p = 0.$$

Therefore the columns  $C_1, C_2, \dots, C_p$  are linearly independent combinations of the columns  $D_{p+1}, \dots, D_n$ . Therefore by 4. and 2., we can, without changing the matrix  $C$ , take the columns  $C_1, \dots, C_p$  instead of  $D_{p+1}, \dots, D_{2p}$  and zeros instead of  $D_{2p+q+1}, \dots, D_n$ .

Then the matrix  $D$  assumes the form

$$D = (D_1, \dots, D_p; C_1, C_2, \dots, C_p; D_{2p+1}, \dots, D_{2p+q}; 0, 0, \dots, 0). \quad (89)$$

In the same way, preserving the forms (87) and (89) of the matrices  $C$  and  $D$ , we can represent the next matrix  $F$  in the form

$$F = (F_1, \dots, F_p; D_1, \dots, D_p; F_{2p+1}, F_{2p+q}; C_1, \dots, C_p; D_{2p+1}, \dots, D_{2p+q}, F_{3p+2q+1}, \dots, F_{3p+2q+r}, 0, \dots, 0), \quad (90)$$

etc.

Formulas (73) gives us the Jordan chains

$$\left. \begin{array}{l} \overbrace{(C_1, D_1, \dots, K_1)}^m, \dots, \overbrace{(C_p, D_p, \dots, K_p)}^m; \\ \underbrace{(D_{2p+1}, F_{2p+1}, \dots, K_{2p+1})}_{m-1}, \dots, \underbrace{(D_{2p+q}, F_{2p+q}, \dots, K_{2p+q})}_{m-1}; \dots \end{array} \right\} \quad (91)$$

These Jordan chains are linearly independent. For all the columns  $C_i$  in (91) are linearly independent, because they form  $p$  linearly independent columns of  $C$ . All the columns  $C_i, D_j$  in (91) are independent, because they form  $2p + q$  independent columns in  $D$ , etc.; finally, all the columns in (91)

are independent, because they form  $n_0 = mp + (m-1)q + \dots$  independent columns in  $K$ . The number of columns in (91) is equal to the sum of the exponents of the elementary divisors corresponding to the given characteristic value  $\lambda_0$ .

Suppose that the matrix  $A = \| a_{ik} \|_1^n$  has  $s$  distinct characteristic values  $\lambda_j$

$$(j = 1, 2, \dots, s;$$

$$\Delta(\lambda) = (\lambda - \lambda_1)^{n_1} (\lambda - \lambda_2)^{n_2} \dots (\lambda - \lambda_s)^{n_s}$$

$$\psi(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_s)^{m_s}.$$

For each characteristic value  $\lambda_j$  we form its system of independent Jordan chains (91); the number of columns in this system is equal to  $n_j$  ( $j = 1, 2, 3, \dots, s$ ). All the chains so obtained contain  $n = n_1 + n_2 + \dots + n_s$  columns.

These  $n$  columns are linearly independent and form one of the required transforming matrices  $T$ .

The proof of the linear independence of these  $n$  columns proceeds as follows.

Every linear combination of these  $n$  columns can be represented in the form

$$\sum_{j=1}^s H_j = 0, \quad (92)$$

where  $H_j$  is a linear combination of columns in the Jordan chains (91) corresponding to the characteristic value  $\lambda_j$  ( $j = 1, 2, \dots, s$ ). But every column in the Jordan chain corresponding to the characteristic value  $\lambda_j$  satisfies the equation

$$(A - \lambda_j E)^{m_j} x = 0.$$

Therefore

$$(A - \lambda_j E)^{m_j} H_j = 0. \quad (93)$$

We take a fixed number  $j$  ( $1 \leq j \leq s$ ) and construct the Lagrange-Sylvester interpolation polynomial  $r(\lambda)$  (See Chapter V, §§ 1, 2) with the following values on the spectrum of the matrix:

$$r(\lambda_i) = r'(\lambda_i) = \dots = r^{(m_i-1)}(\lambda_i) = 0 \text{ for } i \neq j$$

and

$$r(\lambda_j) = 1, r'(\lambda_j) = \dots = r^{(m_j-1)}(\lambda_j) = 0.$$

Then, for every  $i \neq j$ ,  $r(\lambda)$  is divisible by  $(\lambda - \lambda_i)^{m_i}$  without remainder; therefore by (93),

$$r(A) H_i = 0 \quad (i \neq j). \quad (94)$$

In exactly the same way, the difference  $r(\lambda) - 1$  is divisible by  $(\lambda - \lambda_j)^m$  without remainder; therefore

$$r(A)H_j = H_j. \quad (95)$$

Multiplying both sides of (92) by  $r(A)$ , we obtain from (94) and (95):

$$H_j = 0.$$

This is valid for every  $j = 1, 2, \dots, s$ . But  $H_j$  is a linear combination of independent columns corresponding to one and the same characteristic value  $\lambda_j$  ( $j = 1, 2, \dots, s$ ). Therefore all the coefficients in the linear combination  $H_j$  ( $j = 1, 2, \dots, s$ ), and hence all the coefficients in (92), are equal to zero.

*Note.* Let us point out some transformations on the columns of the matrix  $T$  under which it is transformed into the same Jordan form (with the same arrangement of the Jordan diagonal blocks):

I. *Multiplication of all the columns of an arbitrary Jordan chain by a non-zero number.*

II. *Addition to each column (beginning with the second) of a Jordan chain of the preceding column of the same chain, multiplied by one and the same arbitrary number.*

III. *Addition to all the columns of a Jordan chain of the corresponding columns of another chain containing the same or a larger number of columns and corresponding to the same characteristic value.*

*Example 1.*

$$A = \begin{vmatrix} 1 & 0 & 0 & 1 & -1 \\ 0 & 1 & -2 & 3 & -3 \\ 0 & 0 & -1 & 2 & -2 \\ 1 & -1 & 1 & 0 & 1 \\ 1 & -1 & 1 & -1 & 2 \end{vmatrix}; \quad \begin{aligned} \Delta(\lambda) &= (\lambda - 1)^4(\lambda + 1), \\ \psi(\lambda) &= (\lambda - 1)^2(\lambda + 1) = \lambda^3 - \lambda^2 - \lambda + 1. \\ \text{Elementary divisors of the matrix } A & \\ &(\lambda - 1)^2, (\lambda - 1)^2, \lambda + 1. \end{aligned}$$

$$J = \begin{vmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{vmatrix},$$

$$\Psi(\lambda, \mu) = \frac{\psi(\mu) - \psi(\lambda)}{\mu - \lambda} = \mu^2 + (\lambda - 1)\mu + \lambda^2 - \lambda - 1,$$

$$C(\lambda) = \Psi(\lambda E, A) = A^2 + (\lambda - 1)A + (\lambda^2 - \lambda - 1)E.$$

Let us compute successively the column of  $A^2$  and the corresponding columns of  $C(\lambda)$ ,  $C(1)$ ,  $C'(\lambda)$ ,  $C'(1)$ ,  $C'(-1)$ . We must obtain two linearly independent columns of  $C(1)$  and one non-zero column of  $C(-1)$ .

$$\begin{aligned}
C(\lambda) &= \begin{vmatrix} 1 & 0 & 0 & 2* \\ 0 & 1 & 0 & 2* \\ 0 & 0 & 1 & 0* \\ 2 & -2 & 2 & -1* \\ 2 & -2 & 2 & -2* \end{vmatrix} + (\lambda - 1) \begin{vmatrix} 1 & 0 & 0 & 1* \\ 0 & 1 & -2 & 3* \\ 0 & 0 & -1 & 2* \\ 1 & -1 & 1 & 0* \\ 1 & -1 & 1 & -1* \end{vmatrix} + (\lambda^2 - \lambda - 1) \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix} \\
C(+1) &= \begin{vmatrix} 0 & 0 & 0 & 2* \\ 0 & 0 & 0 & 2* \\ 0 & 0 & 0 & 0* \\ 2 & -2 & 2 & -2* \\ 2 & -2 & 2 & -2* \end{vmatrix}, \quad C'(\lambda) = \begin{vmatrix} 1 & 0 & 0 & 1* \\ 0 & 1 & -2 & 3* \\ 0 & 0 & -1 & 2* \\ 1 & -1 & 1 & 0* \\ 1 & -1 & 1 & -1* \end{vmatrix} + (2\lambda - 1) \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix}, \\
C'(+1) &= \begin{vmatrix} 2 & * & * & 1 & * \\ 0 & * & * & 3 & * \\ 0 & * & * & 2 & * \\ 1 & * & * & 1 & * \\ 1 & * & * & -1 & * \end{vmatrix}, \quad C(-1) = \begin{vmatrix} 0 & 0 & 0 & * & * \\ 0 & 0 & 4 & * & * \\ 0 & 0 & 4 & * & * \\ 0 & 0 & 0 & * & * \\ 0 & 0 & 0 & * & * \end{vmatrix}.
\end{aligned}$$

Therefore<sup>38</sup>

$$T = (C_1(+1), C'_1(+1), C_4(+1), C'_4(+1), C_3(-1)) = \begin{vmatrix} 0 & 2 & 2 & 1 & 0 \\ 0 & 0 & 2 & 3 & 4 \\ 0 & 0 & 0 & 2 & 4 \\ 2 & 1 & -2 & 1 & 0 \\ 2 & 1 & -2 & -1 & 0 \end{vmatrix}.$$

The matrix  $T$  can be simplified a little. We

- 1) Divide the fifth column by 4;
- 2) Add the first column to the third and the second to the fourth;
- 3) Subtract the third column from the fourth;
- 4) Divide the first and second columns by 2;
- 5) Subtract the first column, multiplied by  $\frac{1}{2}$ , from the second.

Then we obtain the matrix

$$T_1 = \begin{vmatrix} 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 2 & 1 \\ 1 & 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{vmatrix}.$$

We leave it to the reader to verify that  $AT_1 = T_1J$  and  $|T_1| \neq 0$ .

*Example 2.*

$$A = \begin{vmatrix} 1 & -1 & 1 & -1 \\ -3 & 3 & -5 & 4 \\ 8 & -4 & 3 & -4 \\ 15 & -10 & 11 & -11 \end{vmatrix}; \quad \begin{aligned} \Delta(\lambda) &= (\lambda + 1)^4, \\ \psi(\lambda) &= (\lambda + 1)^3. \end{aligned} \quad \text{Elementary divisors: } (\lambda + 1)^3, \lambda + 1.$$

<sup>38</sup> Here the subscript denotes the number of the column; for example,  $C_3(-1)$  denotes the third column of  $C(-1)$ .



$$J = \begin{vmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix}.$$

We form the polynomials

$$h_1(\lambda) = \frac{\psi(\lambda)}{\lambda+1} = (\lambda+1)^2, \quad h_2(\lambda) = \frac{\psi(\lambda)}{(\lambda+1)^2} = \lambda+1, \quad h_3(\lambda) = \frac{\psi(\lambda)}{(\lambda+1)^3} = 1$$

and the matrices<sup>39</sup>

$$C = h_1(A) = (A + E)^2, \quad D = h_2(A) = A + E, \quad F = E:$$

$$C = \begin{vmatrix} 0 & 0 & 0 & 0 \\ 2 & -1 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ -2 & 1 & -1 & 1 \end{vmatrix}, \quad D = \begin{vmatrix} 2 & -1 & 1 & -1 \\ -3 & 4 & -5 & 4 \\ 8 & -4 & 4 & -4 \\ 15 & -10 & 11 & -10 \end{vmatrix}, \quad F = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}.$$

For the first three columns of  $T$  we take the third column of these matrices:  $T = (C_3, D_3, F_3, *)$ . In the matrices  $C, D, F$ , we subtract twice the third column from the first and we add the third column to the second and to the fourth. We obtain

$$\tilde{C} = \begin{vmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{vmatrix}, \quad \tilde{D} = \begin{vmatrix} 0 & 0 & 1 & 0 \\ 7 & -1 & -5 & -1 \\ 0 & 0 & 4 & 0 \\ -7 & 1 & 11 & 1 \end{vmatrix}, \quad \tilde{F} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -2 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{vmatrix}.$$

In the matrices  $\tilde{D}, \tilde{F}$ , we add the fourth column, multiplied by 7, to the first and subtract the fourth column from the second. We obtain

$$\tilde{C} = \begin{vmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{vmatrix}, \quad \tilde{D} = \begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & -5 & -1 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 11 & 1 \end{vmatrix}, \quad \tilde{F} = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 5 & 0 & 1 & 1 \\ 7 & -1 & 0 & 1 \end{vmatrix}.$$

For the last column of  $T$  we take the first column of  $\tilde{F}$ . Then we have

$$T = (C_3, D_3, F_3, \tilde{F}_1) = \begin{vmatrix} 0 & 1 & 0 & 1 \\ 1 & -5 & 0 & 0 \\ 0 & 4 & 1 & 5 \\ -1 & 11 & 0 & 7 \end{vmatrix}.$$

As a check, we can verify that  $AT = TJ$  and that  $|T| \neq 0$ .

## CHAPTER VII

### THE STRUCTURE OF A LINEAR OPERATOR IN AN $n$ -DIMENSIONAL SPACE

#### *(Geometrical Theory of Elementary Divisors)*

The analytic theory of elementary divisors expounded in the preceding chapter has enabled us to determine for every square matrix a similar matrix having 'normal' or 'canonical' form. On the other hand, we have seen in Chapter III that the behaviour of a linear operator in an  $n$ -dimensional space with respect to various bases is given by means of a class of similar matrices. The existence of a matrix of normal form in such a class is closely connected with important and deep properties of a linear operator in an  $n$ -dimensional space. The study of these properties is the object of the present chapter. The investigation of the structure of a linear operator will lead us, independently of the contents of the preceding chapter, to the theory of transformations of a matrix to a normal form. Therefore the contents of this chapter may be called *the geometrical theory of elementary divisors*.<sup>1</sup>

#### § 1. The Minimal Polynomial of a Vector and a Space (with Respect to a Given Linear Operator)

1. We consider an  $n$ -dimensional vector space  $\mathbf{R}$  over the field  $\mathbf{F}$  and a linear operator  $\mathbf{A}$  in this space.

Let  $\mathbf{x}$  be an arbitrary vector of  $\mathbf{R}$ . We form the sequence of vectors

$$\mathbf{x}, \mathbf{Ax}, \mathbf{A}^2\mathbf{x}, \dots \quad (1)$$

Since the space is finite-dimensional, there is an integer  $p$  ( $0 \leq p \leq n$ ) such that the vectors  $\mathbf{x}, \mathbf{Ax}, \dots, \mathbf{A}^{p-1}\mathbf{x}$  are linearly independent, while  $\mathbf{A}^p\mathbf{x}$  is a linear combination of these vectors with coefficients in  $\mathbf{F}$ :

---

<sup>1</sup> The account of the geometric theory of elementary divisors to be given here is based on our paper [167]. For other geometrical constructions of the theory of elementary divisors, see [22], §§ 96-99 and also [53].

$$A^p \mathbf{x} = -\gamma_1 A^{p-1} \mathbf{x} - \gamma_2 A^{p-2} \mathbf{x} - \dots - \gamma_p \mathbf{x}. \quad (2)$$

We form the monic polynomial  $\varphi(\lambda) = \lambda^p + \gamma_1 \lambda^{p-1} + \dots + \gamma_{p-1} \lambda + \gamma_p$ . (A *monic* polynomial is a polynomial in which the coefficient of the highest power of the variable is unity.) Then (2) can be written:

$$\varphi(A) \mathbf{x} = \mathbf{o}. \quad (3)$$

Every polynomial  $\varphi(\lambda)$  for which (3) holds will be called an *annihilating polynomial for the vector  $\mathbf{x}$* .<sup>2</sup> But it is easy to see that of all the monic annihilating polynomials of  $\mathbf{x}$  the one we have constructed is of least degree. This polynomial will be called the *minimal annihilating polynomial of  $\mathbf{x}$*  or simply the *minimal polynomial of  $\mathbf{x}$* .

Note that every annihilating polynomial  $\tilde{\varphi}(\lambda)$  of  $\mathbf{x}$  is divisible by the minimal polynomial  $\varphi(\lambda)$ .

For let

$$\tilde{\varphi}(\lambda) = \varphi(\lambda) \kappa(\lambda) + \varrho(\lambda),$$

where  $\kappa(\lambda)$ ,  $\varrho(\lambda)$  are quotient and remainder on dividing  $\tilde{\varphi}(\lambda)$  by  $\varphi(\lambda)$ . Then

$$\tilde{\varphi}(A) \mathbf{x} = \kappa(A) \varphi(A) \mathbf{x} + \varrho(A) \mathbf{x} = \varrho(A) \mathbf{x}$$

and therefore  $\varrho(A) \mathbf{x} = \mathbf{o}$ . But the degree of  $\varrho(\lambda)$  is less than that of the minimal polynomial  $\varphi(\lambda)$ . Hence  $\varrho(\lambda) \equiv 0$ .

From what we have proved it follows, in particular, that every vector  $\mathbf{x}$  has only one minimal polynomial.

2. We choose a basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  in  $\mathbf{R}$ . We denote by  $\varphi_1(\lambda), \varphi_2(\lambda), \dots, \varphi_n(\lambda)$  the minimal polynomials of the basis vectors  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  and by  $\psi(\lambda)$  the least common multiple of these polynomials ( $\psi(\lambda)$  is taken with highest coefficient 1). Then  $\psi(\lambda)$  is an annihilating polynomial for all the basis vectors  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ . Since every vector  $\mathbf{x} \in \mathbf{R}$  is representable in the form  $\mathbf{x} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \dots + x_n \mathbf{e}_n$ , we have

$$\psi(A) \mathbf{x} = x_1 \psi(A) \mathbf{e}_1 + x_2 \psi(A) \mathbf{e}_2 + \dots + x_n \psi(A) \mathbf{e}_n = \mathbf{o},$$

i.e.,

$$\psi(A) = \mathbf{O}. \quad (4)$$

The polynomial  $\psi(\lambda)$  is called an *annihilating polynomial for the whole-space  $\mathbf{R}$* . Let  $\tilde{\varphi}(\lambda)$  be an arbitrary annihilating polynomial for the whole space  $\mathbf{R}$ . Then  $\tilde{\varphi}(\lambda)$  is an annihilating polynomial for the basis vectors

<sup>2</sup> Of course, the phrase 'with respect to the given operator  $A$ ' is tacitly understood. For the sake of brevity, this circumstance is not mentioned in the definition, because throughout this entire chapter we shall deal with a single operator  $A$ .

$\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ . Therefore  $\tilde{\psi}(\lambda)$  must be a common multiple of the minimal polynomials  $\varphi_1(\lambda), \varphi_2(\lambda), \dots, \varphi_n(\lambda)$  of these vectors and must therefore be divisible without remainder by their *least common multiple*  $\psi(\lambda)$ . Hence it follows that, of all the annihilating polynomials for the whole space  $\mathbf{R}$ , the one we have constructed,  $\psi(\lambda)$ , has the least degree and it is *monic*. This polynomial is uniquely determined by the space  $\mathbf{R}$  and the operator  $\mathbf{A}$  and is called the *minimal polynomial of the space*  $\mathbf{R}$ .<sup>3</sup> The uniqueness of the minimal polynomial of the space  $\mathbf{R}$  follows from the statement proved above: *every annihilating polynomial  $\tilde{\psi}(\lambda)$  of the space  $\mathbf{R}$  is divisible by the minimal polynomial  $\psi(\lambda)$* . Although the construction of the minimal polynomial  $\psi(\lambda)$  was associated with a definite basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ , the polynomial  $\psi(\lambda)$  itself does not depend on the choice of this basis (this follows from the uniqueness of the minimal polynomial for the space  $\mathbf{R}$ ).

Finally we mention that the minimal polynomial of the space  $\mathbf{R}$  annihilates every vector  $\mathbf{x}$  of  $\mathbf{R}$  so that *the minimal polynomial of the space is divisible by the minimal polynomial of every vector in the space*.

## § 2. Decomposition into Invariant Subspaces with Co-Prime Minimal Polynomials

1. If some collection of vectors  $\mathbf{R}'$  forming part of  $\mathbf{R}$  has the property that the sum of any two vectors of  $\mathbf{R}'$  and the product of any vector of  $\mathbf{R}'$  by a number  $\alpha \in \mathbb{F}$  always belongs to  $\mathbf{R}'$ , then that manifold  $\mathbf{R}'$  is itself a vector space, a *subspace* of  $\mathbf{R}$ .

If two subspaces  $\mathbf{R}'$  and  $\mathbf{R}''$  of  $\mathbf{R}$  are given and if it is known that

1.  $\mathbf{R}'$  and  $\mathbf{R}''$  have no vector in common except the null vector, and
2. every vector  $\mathbf{x}$  of  $\mathbf{R}$  can be represented in the form of a sum

$$\mathbf{x} = \mathbf{x}' + \mathbf{x}'' \quad (\mathbf{x}' \in \mathbf{R}', \mathbf{x}'' \in \mathbf{R}''), \quad (5)$$

then we shall say that the space  $\mathbf{R}$  is *decomposed* into the two subspaces  $\mathbf{R}'$  and  $\mathbf{R}''$  and shall write:

$$\mathbf{R} = \mathbf{R}' + \mathbf{R}'' \quad (6)$$

Note that the condition 1. implies the uniqueness of the representation (5). For if for a certain vector  $\mathbf{x}$  we had two distinct representations in the form of a sum of terms from  $\mathbf{R}'$  and  $\mathbf{R}''$ , (5) and

$$\mathbf{x} = \tilde{\mathbf{x}}' + \tilde{\mathbf{x}}'' \quad (\tilde{\mathbf{x}}' \in \mathbf{R}', \tilde{\mathbf{x}}'' \in \mathbf{R}'') \quad (7)$$

then, subtracting (7) from (5) term by term, we would obtain:

<sup>3</sup> If in some basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  a matrix  $\mathbf{A} = \|a_{ik}\|$  then the annihilating or minimal polynomial of the space  $\mathbf{R}$  (with respect to  $\mathbf{A}$ ) is the annihilating or minimal polynomial of the matrix  $\mathbf{A}$ , and vice versa. Compare with Chapter IV, § 6.

$$\mathbf{x}' - \tilde{\mathbf{x}}' = \tilde{\mathbf{x}}'' - \mathbf{x}'',$$

i.e., equality of the non-null vectors  $\mathbf{x}' - \tilde{\mathbf{x}}' \in \mathbf{R}'$  and  $\tilde{\mathbf{x}}'' - \mathbf{x}'' \in \mathbf{R}''$ , which, by 1., is impossible.

Thus, condition 1. may be replaced by the requirement that the representation (5) be unique. In this form, the definition of decomposition immediately extends to an arbitrary number of subspaces.

Let

$$\mathbf{R} = \mathbf{R}' + \mathbf{R}''$$

and let  $\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_{n'}$  and  $\mathbf{e}''_1, \mathbf{e}''_2, \dots, \mathbf{e}''_{n''}$  be bases of  $\mathbf{R}'$  and  $\mathbf{R}''$ , respectively. Then the reader can easily prove that all these  $n' + n''$  vectors are linearly independent and form a basis of  $\mathbf{R}$ , so that a basis of the whole space is formed from bases of the subspaces. It follows, in particular, that  $n = n' + n''$ .

*Example 1.* Suppose that in a three-dimensional space three directions, not parallel to one another and not in the same plane, are given. Since every vector in the space can be split, uniquely, into components in these three directions, we have

$$\mathbf{R} = \mathbf{R}' + \mathbf{R}'' + \mathbf{R}''',$$

where  $\mathbf{R}$  is the set of all the vectors of our space,  $\mathbf{R}'$  the set of all vectors parallel to the first direction,  $\mathbf{R}''$  to the second, and  $\mathbf{R}'''$  to the third. In this case,  $n = 3$  and  $n' = n'' = n''' = 1$ .

*Example 2.* Suppose that in a three-dimensional space a plane and a line intersecting the plane are given. Then

$$\mathbf{R} = \mathbf{R}' + \mathbf{R}'',$$

where  $\mathbf{R}$  is the set of all vectors of our space,  $\mathbf{R}'$  the set of all vectors parallel to the given plane, and  $\mathbf{R}''$  the set of all vectors parallel to the given line. In this example,  $n = 3$ ,  $n' = 2$ ,  $n'' = 1$ .

**2.** A subspace  $\mathbf{R}' \subset \mathbf{R}$  is called *invariant* with respect to the operator  $\mathbf{A}$  if  $\mathbf{A}\mathbf{R}' \subset \mathbf{R}'$ , i.e. if  $\mathbf{x} \in \mathbf{R}'$  implies  $\mathbf{A}\mathbf{x} \in \mathbf{R}'$ . In other words, the operator  $\mathbf{A}$  carries a vector of an invariant subspace into a vector of the same subspace.

In what follows we shall carry out a decomposition of the whole space into subspaces invariant with respect to  $\mathbf{A}$ . The decomposition reduces the study of the behavior of an operator in the whole space to the study of its behavior in the various component subspaces.

We shall now prove the following theorem:

**THEOREM 1** (First Theorem on the Decomposition of a Space into Invariant Subspaces): *If for a given operator  $A$  the minimal polynomial  $\psi(\lambda)$  of the space is represented over  $\mathbb{F}$  in the form of a product of two co-prime polynomials  $\psi_1(\lambda)$  and  $\psi_2(\lambda)$  (with highest coefficients 1)*

$$\psi(\lambda) = \psi_1(\lambda) \psi_2(\lambda), \quad (8)$$

*then the whole space  $\mathbf{R}$  splits into two invariant subspaces  $\mathbf{I}_1$  and  $\mathbf{I}_2$*

$$\mathbf{R} = \mathbf{I}_1 + \mathbf{I}_2, \quad (9)$$

*whose minimal polynomials are  $\psi_1(\lambda)$  and  $\psi_2(\lambda)$ , respectively.*

*Proof.* We denote by  $\mathbf{I}_1$  the set of all vectors  $\mathbf{x} \in \mathbf{R}$  satisfying the equation  $\psi_1(A) \mathbf{x} = \mathbf{o}$ .  $\mathbf{I}_2$  is similarly defined by the equation  $\psi_2(A) \mathbf{x} = \mathbf{o}$ .  $\mathbf{I}_1$  and  $\mathbf{I}_2$  so defined are subspaces of  $\mathbf{R}$ .

Since  $\psi_1(\lambda)$  and  $\psi_2(\lambda)$  are co-prime, it follows that there exist polynomials  $\chi_1(\lambda)$  and  $\chi_2(\lambda)$  (with coefficients in  $\mathbb{F}$ ) such that

$$1 = \psi_1(\lambda) \chi_1(\lambda) + \psi_2(\lambda) \chi_2(\lambda). \quad (10)$$

Now let  $\mathbf{x}$  be an arbitrary vector of  $\mathbf{R}$ . In (10) we replace  $\lambda$  by  $A$  and we apply both sides of the operator equation so obtained to the vector  $\mathbf{x}$ :

$$\mathbf{x} = \psi_1(A) \chi_1(A) \mathbf{x} + \psi_2(A) \chi_2(A) \mathbf{x}, \quad (11)$$

i.e.,

$$\mathbf{x} = \mathbf{x}' + \mathbf{x}'', \quad (12)$$

where

$$\mathbf{x}' = \psi_2(A) \chi_2(A) \mathbf{x}, \quad \mathbf{x}'' = \psi_1(A) \chi_1(A) \mathbf{x}. \quad (13)$$

Furthermore,

$$\psi_1(A) \mathbf{x}' = \psi(A) \chi_2(A) \mathbf{x} = \mathbf{o}, \quad \psi_2(A) \mathbf{x}'' = \psi(A) \chi_1(A) \mathbf{x} = \mathbf{o},$$

i.e.,

$$\mathbf{x}' \in \mathbf{I}_1 \text{ and } \mathbf{x}'' \in \mathbf{I}_2.$$

$\mathbf{I}_1$  and  $\mathbf{I}_2$  have only the null vector in common. For if  $\mathbf{x}_0 \in \mathbf{I}_1$  and  $\mathbf{x}_0 \in \mathbf{I}_2$ , i.e.,  $\psi_1(A) \mathbf{x}_0 = \mathbf{o}$  and  $\psi_2(A) \mathbf{x}_0 = \mathbf{o}$ , then by (11)

$$\mathbf{x}_0 = \chi_1(A) \psi_1(A) \mathbf{x}_0 + \chi_2(A) \psi_2(A) \mathbf{x}_0 = \mathbf{o}.$$

Thus we have proved that  $\mathbf{R} = \mathbf{I}_1 + \mathbf{I}_2$ .

Now suppose that  $\mathbf{x} \in \mathbf{I}_1$ . Then  $\psi_1(\mathbf{A})\mathbf{x} = \mathbf{o}$ . Multiplying both sides of this equation by  $\mathbf{A}$  and reversing the order of  $\mathbf{A}$  and  $\psi_1(\mathbf{A})$ , we obtain  $\psi_1(\mathbf{A})\mathbf{A}\mathbf{x} = \mathbf{o}$ , i.e.,  $\mathbf{A}\mathbf{x} \in \mathbf{I}_1$ . This proves that the subspace  $\mathbf{I}_1$  is invariant with respect to  $\mathbf{A}$ . The invariance of the subspace  $\mathbf{I}_2$  is proved similarly.

We shall now show that  $\psi_1(\lambda)$  is the minimal polynomial of  $\mathbf{I}_1$ . Let  $\tilde{\psi}_1(\lambda)$  be an arbitrary annihilating polynomial for  $\mathbf{I}_1$ , and  $\mathbf{x}$  an arbitrary vector of  $\mathbf{R}$ . Using the decomposition (12) already established, we write:

$$\tilde{\psi}_1(\mathbf{A})\psi_2(\mathbf{A})\mathbf{x} = \psi_2(\mathbf{A})\tilde{\psi}_1(\mathbf{A})\mathbf{x}' + \tilde{\psi}_1(\mathbf{A})\psi_2(\mathbf{A})\mathbf{x}'' = \mathbf{o}.$$

Since  $\mathbf{x}$  is an arbitrary vector of  $\mathbf{R}$ , it follows that the product  $\tilde{\psi}_1(\lambda)\psi_2(\lambda)$  is an annihilating polynomial for  $\mathbf{R}$  and is therefore divisible by  $\psi(\lambda) = \psi_1(\lambda)\psi_2(\lambda)$  without remainder; in other words,  $\psi_1(\lambda)$  is divisible by  $\tilde{\psi}_1(\lambda)$ . But  $\tilde{\psi}_1(\lambda)$  is an arbitrary annihilating polynomial for  $\mathbf{I}_1$  and  $\psi_1(\lambda)$  is a particular one of the annihilating polynomials (by the definition of  $\mathbf{I}_1$ ). Hence  $\psi_1(\lambda)$  is the minimal polynomial of  $\mathbf{I}_1$ . In exactly the same way it is shown that  $\psi_2(\lambda)$  is the minimal polynomial for the invariant subspace  $\mathbf{I}_2$ .

This completes the proof of the theorem.

Let us decompose  $\psi(\lambda)$  into irreducible factors over  $\mathbb{F}$ :

$$\psi(\lambda) = [\varphi_1(\lambda)]^{e_1} [\varphi_2(\lambda)]^{e_2} \cdots [\varphi_s(\lambda)]^{e_s} \quad (14)$$

(here  $\varphi_1(\lambda), \varphi_2(\lambda), \dots, \varphi_s(\lambda)$  are distinct irreducible polynomials over  $\mathbb{F}$  with highest coefficient 1). Then by the theorem we have

$$\mathbf{R} = \mathbf{I}_1 + \mathbf{I}_2 + \dots + \mathbf{I}_s, \quad (15)$$

where  $\mathbf{I}_k$  is an invariant subspace with the minimal polynomial  $[\varphi_k(\lambda)]^{e_k}$  ( $k = 1, 2, \dots, s$ )

Thus, the theorem reduces the study of the behaviour of a linear operator in an arbitrary space to the study of the behaviour of this operator in a space where the minimal polynomial is a power of an irreducible polynomial over  $\mathbb{F}$ . We shall take advantage of this to prove the following important theorem:

**THEOREM 2:** *In a vector space there always exists a vector whose minimal polynomial coincides with the minimal polynomial of the whole space.*

We consider first the special case where the minimal polynomial of the space  $\mathbf{R}$  is a power of an irreducible polynomial  $\varphi(\lambda)$ :

$$\psi(\lambda) = [\varphi(\lambda)]^i.$$

In  $R$  we choose a basis  $e_1, e_2, \dots, e_n$ . The minimal polynomial of  $e_1$  is a divisor of  $\psi(\lambda)$  and is therefore representable in the form  $[\varphi(\lambda)]^{l_1}$ , where  $l_1 \leq l$  ( $i=1, 2, \dots, n$ ).

But the minimal polynomial of the space is the least common multiple of the minimal polynomials of the basis vectors, so that  $\psi(\lambda)$  is the largest of the powers  $[\varphi(\lambda)]^{l_i}$  ( $i=1, 2, \dots, n$ ). In other words,  $\psi(\lambda)$  coincides with the minimal polynomial of one of the basis vectors  $e_1, e_2, \dots, e_n$ .

Turning now to the general case, we prove the following preliminary lemma:

**LEMMA:** *If the minimal polynomials of the vectors  $e'$  and  $e''$  are co-prime, then the minimal polynomial of the sum vector  $e' + e''$  is equal to the product of the minimal polynomials of the constituent vectors.*

*Proof.* Let  $\chi_1(\lambda)$  and  $\chi_2(\lambda)$  be the minimal polynomials of the vectors  $e'$  and  $e''$ . By assumption,  $\chi_1(\lambda)$  and  $\chi_2(\lambda)$  are co-prime. Let  $\chi(\lambda)$  be an arbitrary annihilating polynomial of the vector  $e = e' + e''$ . Then

$$\chi_2(A) \chi(A) e' = \chi_2(A) \chi(A) e - \chi(A) \chi_2(A) e'' = 0,$$

i.e.,  $\chi_2(\lambda) \chi(\lambda)$  is an annihilating polynomial of  $e'$ . Therefore  $\chi_2(\lambda) \chi(\lambda)$  is divisible by  $\chi_1(\lambda)$ , and since  $\chi_1(\lambda)$  and  $\chi_2(\lambda)$  are co-prime,  $\chi(\lambda)$  is divisible by  $\chi_1(\lambda)$ . It is proved similarly that  $\chi(\lambda)$  is divisible by  $\chi_2(\lambda)$ . But  $\chi_1(\lambda)$  and  $\chi_2(\lambda)$  are co-prime. Therefore  $\chi(\lambda)$  is divisible by the product  $\chi_1(\lambda) \chi_2(\lambda)$ . Thus, every annihilating polynomial of the vector  $e$  is divisible by  $\chi_1(\lambda) \chi_2(\lambda)$ . Therefore  $\chi_1(\lambda) \chi_2(\lambda)$  is the minimal polynomial of the vector  $e = e' + e''$ .

We now return to Theorem 2. For the proof in the general case we use the decomposition (15). Since the minimal polynomials of the subspaces  $I_1, I_2, \dots, I_s$  are powers of irreducible polynomials, our assertion is already proved for these subspaces. Therefore there exist vectors  $e' \in I_1, e'' \in I_2, \dots, e^{(s)} \in I_s$  whose minimal polynomials are  $[\varphi_1(\lambda)]^{r_1}, [\varphi_2(\lambda)]^{r_2}, \dots, [\varphi_s(\lambda)]^{r_s}$ , respectively. By the lemma, the minimal polynomial of the vector  $e = e' + e'' + \dots + e^{(s)}$  is equal to the product

$$[\varphi_1(\lambda)]^{r_1} [\varphi_2(\lambda)]^{r_2} \dots [\varphi_s(\lambda)]^{r_s},$$

i.e., to the minimal polynomial of the space  $R$ .

### § 3. Congruence. Factor Space

**L.** Suppose given a subspace  $I \subset R$ . We shall say that two vectors  $x, y$  of  $R$  are *congruent modulo  $I$*  and shall write  $x \equiv y \pmod{I}$  if and only if  $y - x \in I$ . It is easy to verify that the concept of congruence so introduced has the following properties:



For all  $x, y, z \in R$

1.  $x \equiv x \pmod{I}$  (reflexivity of congruence).
2. From  $x \equiv y \pmod{I}$  it follows that  $y \equiv x \pmod{I}$  (symmetry of congruence).
3. From  $x \equiv y \pmod{I}$  and  $y \equiv z \pmod{I}$  it follows that  $x \equiv z \pmod{I}$  (transitivity of congruence).

The presence of these three properties enables us to make use of congruence to divide all the vectors of the space into classes, by assigning vectors that are pairwise congruent  $\pmod{I}$  to the same class (vectors of distinct classes are incongruent  $\pmod{I}$ ). The class containing the vector  $x$  will be denoted by  $\bar{x}$ .<sup>5</sup> The subspace  $I$  is one of these classes, namely  $\bar{0}$ . Note that to every congruence  $x \equiv y \pmod{I}$  there corresponds the equality<sup>6</sup> of the associated classes:  $\bar{x} = \bar{y}$ .

It is elementary to prove that congruences may be added term by term and multiplied by a number of  $F$ :

1. From

$$x \equiv x' \text{ and } y \equiv y' \pmod{I}$$

it follows that

$$x + y \equiv x' + y' \pmod{I}.$$

2. From

$$x \equiv x' \pmod{I}$$

it follows that

$$ax \equiv ax' \pmod{I} \quad (a \in F).$$

These properties of congruence show that the operations of addition and multiplication by a number of  $F$  do not 'break up' the classes. If we take two classes  $\bar{x}$  and  $\bar{y}$  and add elements  $x, x', \dots$  of the first class to arbitrary elements  $y, y', \dots$  of the second class, then all the sums so obtained belong to one and the same class, which we call the *sum* of the classes  $\bar{x}$  and  $\bar{y}$  and denote by  $\bar{x} + \bar{y}$ . Similarly, if all the vectors  $x, x', \dots$  of the class  $\bar{x}$  are multiplied by a number  $a \in F$ , then the products belong to one class, which we denote by  $a\bar{x}$ .

Thus, in the manifold  $\bar{R}$  of all classes  $\bar{x}, \bar{y}, \dots$  two operations are introduced: 'addition' and 'multiplication by a number of  $F$ .' It is easy to verify that these operations have the properties set forth in the definition of a vector space (Chapter III, § 1). Therefore  $\bar{R}$ , as well as  $R$ , is a vector

<sup>5</sup> Since each class contains an infinite set of vectors, there is, by this condition, a infinite number of ways of designating the class.

<sup>6</sup> That is, identity.

space over the field  $\mathbb{F}$ . We shall say that  $\bar{\mathbf{R}}$  is a *factor space* of  $\mathbf{R}$ . If  $n, m, \bar{n}$  are the dimensions of the spaces  $\mathbf{R}, \mathbf{I}, \bar{\mathbf{R}}$ , respectively, then  $\bar{n} = n - m$ .

2. All the concepts introduced in this section can be illustrated very well by the following example.

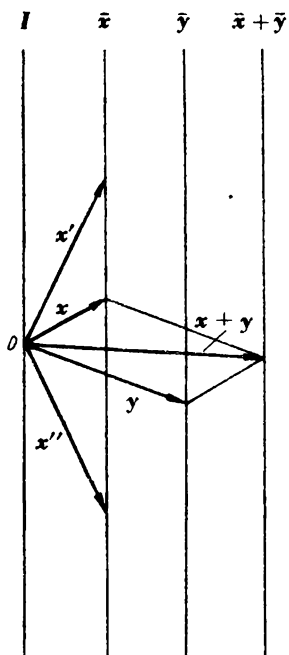


Fig. 4

*Example.* Let  $\mathbf{R}$  be the set of all vectors of a three-dimensional space and  $\mathbb{F}$  the field of real numbers. For greater clarity, we shall represent vectors in the form of directed segments beginning at a point  $O$ . Let  $\mathbf{I}$  be a straight line passing through  $O$  (more accurately: the set of vectors that lie along some line passing through  $O$ ; Fig. 4.)

The congruence  $\mathbf{x} \equiv \mathbf{x}' \pmod{\mathbf{I}}$  signifies that the vectors  $\mathbf{x}$  and  $\mathbf{x}'$  differ by a vector of  $\mathbf{I}$ , i.e., the segment containing the end-points of  $\mathbf{x}$  and  $\mathbf{x}'$  is parallel to  $\mathbf{I}$ . Therefore the class  $\bar{\mathbf{x}}$  is represented by the line passing through the end-point of  $\mathbf{x}$  and parallel to  $\mathbf{I}$  (more accurately: by the 'bundle' of vectors starting from  $O$  whose end-points lie on that line). 'Bundles' may be added and multiplied by a real number (by adding and multiplying the vectors that occur in the bundles). These 'bundles' are also the elements of the factor space  $\bar{\mathbf{R}}$ . In this example,  $n=3$ ,  $m=1$ ,  $\bar{n}=2$ .

We obtain another example by taking for  $\mathbf{I}$  a plane passing through  $O$ . In this example,  $n=3$ ,  $m=2$ ,  $\bar{n}=1$ .

Now let  $\mathbf{A}$  be a linear operator in  $\mathbf{R}$ . Let us assume that  $\mathbf{I}$  is an *invariant subspace with respect to  $\mathbf{A}$* . The reader will easily prove that from  $\mathbf{x} \equiv \mathbf{x}' \pmod{\mathbf{I}}$  it follows that  $\mathbf{Ax} \equiv \mathbf{Ax}' \pmod{\mathbf{I}}$ , so that the operator  $\mathbf{A}$  can be applied to both sides of a congruence. In other words, if the operator  $\mathbf{A}$  is applied to all vectors  $\mathbf{x}, \mathbf{x}', \dots$  of a class  $\bar{\mathbf{x}}$ , then the vectors  $\mathbf{Ax}, \mathbf{Ax}', \dots$  also belong to one class, which we denote by  $\mathbf{A}\bar{\mathbf{x}}$ . The linear operator  $\mathbf{A}$  carries classes into classes and is, thus, a linear operator in  $\bar{\mathbf{R}}$ .

We shall say that the vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$  are *linearly dependent modulo  $\mathbf{I}$*  if there exist numbers  $\alpha_1, \alpha_2, \dots, \alpha_p$  in  $\mathbb{F}$ , not all equal to zero, such that

$$\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \dots + \alpha_p \mathbf{x}_p \equiv \mathbf{0} \pmod{\mathbf{I}}. \quad (16)$$

Note that not only the concept of linear dependence of vectors, but also all the concepts, statements, and reasonings, in the preceding sections of this chapter can be repeated word for word with the symbol '=' replaced throughout by the symbol ' $\equiv (\text{mod } I)$ ,' where  $I$  is some fixed subspace invariant with respect to  $A$ .

Thus, we can introduce the concepts of an annihilating polynomial and of the minimal polynomial of a vector or a space  $(\text{mod } I)$ . All these concepts will be called 'relative,' in contrast to the 'absolute' concepts that were introduced earlier (and that hold for the symbol '=').

The reader should observe that *the relative minimal polynomial (of a vector or a space) is a divisor of the absolute one*. For example, let  $\sigma_1(\lambda)$  be the relative minimal polynomial of a vector  $x$  and  $\sigma(\lambda)$  the corresponding absolute minimal polynomial.

Then

$$\sigma(A)x = o,$$

and hence it follows that also

$$\sigma(A)x \equiv o \pmod{I}.$$

Therefore  $\sigma(\lambda)$  is a relative annihilating polynomial of  $x$  and as such is divisible by the relative minimal polynomial  $\sigma_1(\lambda)$ .

Side by side with the 'absolute' statements of the preceding sections we have 'relative' statements. For example, we have the statement: 'In every space there always exists a vector whose relative minimal polynomial coincides with the relative minimal polynomial of the whole space.'

The truth of all 'relative' statements depends on the fact that by operating with congruences modulo  $I$  we deal essentially with equalities—however not in the space  $R$ , but in the space  $\bar{R}$ .

#### § 4. Decomposition of a Space into Cyclic Invariant Subspaces

1. Let  $\sigma(\lambda) = \lambda^p + \alpha_1 \lambda^{p-1} + \dots + \alpha_{p-1} \lambda + \alpha_p$  be the minimal polynomial of a vector  $e$ . Then the vectors

$$e, Ae, \dots, A^{p-1}e \tag{1}$$

are linearly independent, and

$$A^p e = -\alpha_p e - \alpha_{p-1} A e - \dots - \alpha_1 A^{p-1} e. \tag{1}$$

The vectors (17) form a basis of a  $p$ -dimensional subspace  $I$ . We shall call this subspace *cyclic* in view of the special character of the basis (17) and of (18).<sup>7</sup> The operator  $A$  carries the first vector of (17) into the second, the second into the third, etc. The last basis vector is carried by  $A$  into a linear combination of the basis vectors in accordance with (18). Thus,  $A$  carries every basis vector into a vector of  $I$  and hence an arbitrary vector of  $I$  into another vector of  $I$ . In other words, a *cyclic subspace is always invariant with respect to  $A$* .

Every vector  $x \in I$  is representable in the form of a linear combination of the basis vectors (17), i.e., in the form

$$x = \chi(A) e, \quad (19)$$

where  $\chi(\lambda)$  is a polynomial in  $\lambda$  of degree  $\leq p - 1$  with coefficients in  $F$ . By forming all possible polynomials  $\chi(\lambda)$  of degree  $\leq p - 1$  with coefficients in  $F$  we obtain all the vectors of  $I$ , each once only, i.e., for only one polynomial  $\chi(\lambda)$ . In view of the basis (17) or the formula (19) we shall say that the vector  $e$  *generates* the subspace.

Note that *the minimal polynomial of the generating vector  $e$  is also the minimal polynomial of the whole subspace  $I$* .

2. We are now ready to establish the fundamental proposition of the whole theory, according to which the space  $R$  splits into cyclic subspaces.

Let  $\psi_1(\lambda) = \psi(\lambda) = \lambda^m + \alpha_1 \lambda^{m-1} + \dots + \alpha_m$  be the minimal polynomial of the space  $R$ . Then there exists a vector  $e$  in the space for which this polynomial is minimal (Theorem 2, p. 180). Let  $I_1$  denote the cyclic subspace with the basis

$$e, Ae, \dots, A^{m-1}e. \quad (20)$$

If  $n = m$ , then  $R = I_1$ . Suppose that  $n > m$  and that the polynomial

$$\psi_2(\lambda) = \lambda^p + \beta_1 \lambda^{p-1} + \dots + \beta_p$$

is the minimal polynomial of  $R \pmod{I_1}$ . By the remark at the end of § 3,  $\psi_2(\lambda)$  is a divisor of  $\psi_1(\lambda)$ , i.e., there exists a polynomial  $\kappa(\lambda)$  such that

$$\psi_1(\lambda) = \psi_2(\lambda) \kappa(\lambda). \quad (21)$$

<sup>7</sup> It would be more accurate to call this subspace: cyclic with respect to the linear operator  $A$ . But since the whole theory is built up with reference to a single operator  $A$ , the words 'with respect to the linear operator  $A$ ' are omitted for the sake of brevity (see the similar remark in footnote 2, p. 176).

Moreover, in  $R$  there exists a vector  $g^*$  whose relative minimal polynomial is  $\psi_2(\lambda)$ . Then

$$\psi_2(A)g^* \equiv 0 \pmod{I_1}, \quad (22)$$

i.e., there exists a polynomial  $\chi(\lambda)$  of degree  $\leq m-1$  such that

$$\psi_2(A)g^* = \chi(A)e. \quad (23)$$

We apply the operator  $\kappa(A)$  to both sides of the equation. Then by (21) we obtain on the left  $\psi_1(A)g^*$ , i.e. zero, because  $\psi_1(\lambda)$  is the absolute minimal polynomial of the space; therefore

$$\kappa(A)\chi(A)e = 0.$$

This equation shows that the product  $\kappa(\lambda)\chi(\lambda)$  is an annihilating polynomial of the vector  $e$  and is therefore divisible by the minimal polynomial  $\psi_1(\lambda) = \kappa(\lambda)\psi_2(\lambda)$ , so that  $\chi(\lambda)$  is divisible by  $\psi_2(\lambda)$ :

$$\chi(\lambda) = \kappa_1(\lambda)\psi_2(\lambda), \quad (24)$$

where  $\kappa_1(\lambda)$  is a polynomial. Using this decomposition of  $\chi(\lambda)$ , we may rewrite (23) as follows:

$$\psi_2(A)[g^* - \kappa_1(A)e] = 0. \quad (25)$$

We now introduce the vector

$$g = g^* - \kappa_1(A)e. \quad (26)$$

Then (25) can be written as follows:

$$\psi_2(A)g = 0. \quad (27)$$

The last equation shows that  $\psi_2(\lambda)$  is an absolute annihilating polynomial of the vector  $g$  and is therefore divisible by the absolute minimal polynomial of  $g$ . On the other hand, we have from (26):

$$g \equiv g^* \pmod{I_1}. \quad (28)$$

Hence  $\psi_2(\lambda)$ , being the relative minimal polynomial of  $g^*$ , is the same for  $g$  as well. Comparing the last two statements, we deduce that  $\psi_2(\lambda)$  is simultaneously the relative and the absolute minimal polynomial of  $g$ .

From the fact that  $\psi_2(\lambda)$  is the absolute minimal polynomial of  $g$  it follows that the subspace  $I_2$  with the basis

$$g, Ag, \dots, A^{p-1}g \quad (29)$$

is cyclic.

From the fact that  $\psi_2(\lambda)$  is the relative minimal polynomial of  $g \pmod{I_1}$  it follows that the vectors (29) are linearly independent  $\pmod{I_1}$ , i.e., no linear combination with coefficients not all zero can be equal to a linear combination of the vectors (20). Since the latter are themselves linearly independent, our last statement asserts the linear independence of the  $m + p$  vectors

$$e, Ae, \dots, A^{m-1}e; g, Ag, \dots, A^{p-1}g. \quad (30)$$

The vectors (30) form a basis of the invariant subspace  $I_1 + I_2$  of dimension  $m + p$ .

If  $n = m + p$ , then  $R = I_1 + I_2$ . If  $n > m + p$ , we consider  $R \pmod{I_1 + I_2}$  and continue our process of separating cyclic subspaces. Since the whole space  $R$  is of finite dimension  $n$ , this process must come to an end with some subspace  $I_t$ , where  $t \leq n$ .

We have arrived at the following theorem:

**THEOREM 3** (Second Theorem on the Decomposition of a Space into Invariant Subspaces): *Relative to a given linear operator  $A$  the space can always be split into cyclic subspaces  $I_1, I_2, \dots, I_t$  with minimal polynomials  $\psi_1(\lambda), \psi_2(\lambda), \dots, \psi_t(\lambda)$*

$$R = I_1 + I_2 + \dots + I_t \quad (31)$$

such that  $\psi_1(\lambda)$  coincides with the minimal polynomial  $\psi(\lambda)$  of the whole space and that each  $\psi_i(\lambda)$  is divisible by  $\psi_{i-1}(\lambda)$  ( $i = 2, 3, \dots, t$ ).

**3.** We now mention some properties of cyclic spaces. Let  $R$  be a cyclic  $n$ -dimensional space and  $\psi(\lambda) = \lambda^m + \dots$  its minimal polynomial. Then it follows from the definition of a cyclic space that  $m = n$ . Conversely, suppose that  $R$  is an arbitrary space and that it is known that  $m = n$ . Applying the proof of the decomposition theorem, we represent  $R$  in the form (31). But the dimension of the cyclic subspace  $I_1$  is  $m$ , because its minimal polynomial coincides with the minimal polynomial of the whole space. Since  $m = n$  by assumption, we have  $R = I_1$ , i.e.,  $R$  is a cyclic space.

Thus we have established the following criterion for cyclicity of a space:

**THEOREM 4:** *A space is cyclic if and only if its dimension is equal to the degree of its minimal polynomial.*

Next, suppose that we have a decomposition of a cyclic space  $R$  into two invariant subspaces  $I_1$  and  $I_2$ :

$$R = I_1 + I_2. \quad (32)$$

We denote the dimensions of  $\mathbf{R}$ ,  $\mathbf{I}_1$ , and  $\mathbf{I}_2$  by  $n$ ,  $n_1$ , and  $n_2$ , their minimal polynomials by  $\psi(\lambda)$ ,  $\psi_1(\lambda)$ , and  $\psi_2(\lambda)$ , and the degrees of these minimal polynomials by  $m$ ,  $m_1$ , and  $m_2$ , respectively. Then

$$m_1 \leq n_1, \quad m_2 \leq n_2. \quad (33)$$

We add these inequalities term by term:

$$m_1 + m_2 \leq n_1 + n_2. \quad (34)$$

Since  $\psi(\lambda)$  is the least common multiple of  $\psi_1(\lambda)$  and  $\psi_2(\lambda)$ , we have

$$m \leq m_1 + m_2. \quad (35)$$

Moreover, it follows from (32) that

$$n = n_1 + n_2. \quad (36)$$

(34), (35), and (36) give us a chain of relations

$$m \leq m_1 + m_2 \leq n_1 + n_2 = n. \quad (37)$$

But since the space  $\mathbf{R}$  is cyclic, the extreme numbers of this chain,  $m$  and  $n$ , are equal. Therefore we have equality in the middle terms, i.e.,

$$m = m_1 + m_2 = n_1 + n_2.$$

From the fact that  $m = m_1 + m_2$  we deduce that  $\psi_1(\lambda)$  and  $\psi_2(\lambda)$  are co-prime.

Bearing (33) in mind, we find from  $m_1 + m_2 = n_1 + n_2$  that

$$m_1 = n_1 \quad \text{and} \quad m_2 = n_2. \quad (38)$$

These equations mean that the subspaces  $\mathbf{I}_1$  and  $\mathbf{I}_2$  are cyclic.

Thus we have arrived at the following proposition:

**THEOREM 5:** *A cyclic space can only split into invariant subspaces that 1. are also cyclic and 2. have co-prime minimal polynomials.*

The same arguments (in the opposite order) show that Theorem 5 has converse:

**THEOREM 6:** *If a space is split into invariant subspaces that 1. are cyclic and 2. have co-prime minimal polynomials, then the space itself is cyclic.*

Suppose now that  $\mathbf{R}$  is a cyclic space and that its minimal polynomial is a power of an irreducible polynomial over  $\mathbb{F}$ :  $\psi(\lambda) = [\varphi(\lambda)]^e$ . In this case the minimal polynomial of every invariant subspace of  $\mathbf{R}$  must also be a power of this irreducible polynomial  $\varphi(\lambda)$ . Therefore the minimal polynomials of any two invariant subspaces cannot be co-prime. But then, from what we have proved,  $\mathbf{R}$  cannot split into invariant subspaces.





*Note.* Theorem 8 (the third decomposition theorem) has been proved by applying the first two decomposition theorems. But it can also be obtained by other means, namely, as an immediate (and almost trivial) corollary of Theorem 7.

For if the space  $R$  splits at all, then it can always be split into indecomposable invariant subspaces:

$$R = I' + I'' + \dots + I^{(s)}. \quad (40)$$

By Theorem 7, each of the constituent subspaces is cyclic and has as its minimal polynomial a power of an irreducible polynomial over  $F$ .

### § 5. The Normal Form of a Matrix

1. Let  $I_1$  be an  $m$ -dimensional invariant subspace of  $R$ . In  $I_1$  we take an arbitrary basis  $e_1, e_2, \dots, e_m$  and complement it to a basis

$$e_1, e_2, \dots, e_m, e_{m+1}, \dots, e_n$$

of  $R$ . Let us see what the matrix  $A$  of the operator  $A$  looks like in this basis. We remind the reader that the  $k$ -th column of  $A$  consists of the coordinates of the vector  $Ae_k$  ( $k = 1, 2, \dots, n$ ). For  $k \leq m$  the vector  $Ae_k \in I_1$  (by the invariance of  $I_1$ ) and the last  $n - m$  coordinates of  $Ae_k$  are zero. Therefore  $A$  has the following form

$$A = \left\| \begin{array}{c|c} \overbrace{A_1}^m & \overbrace{A_2}^{n-m} \\ \hline 0 & A_3 \end{array} \right\| \begin{array}{l} \} m \\ \} n-m \end{array}, \quad (41)$$

where  $A_1$  and  $A_2$  are square matrices of orders  $m$  and  $n - m$ , respectively, and  $A_3$  is a rectangular matrix. The fact that the fourth 'block' is zero expresses the invariance of the subspace  $I_1$ . The matrix  $A_1$  gives the operator  $A$  in  $I_1$  (with respect to the basis  $e_1, e_2, \dots, e_m$ ).

Let us assume now that  $e_{m+1}, \dots, e_n$  is the basis of some invariant subspace  $I_2$ , so that  $R = I_1 + I_2$  and a basis of the whole space is formed from the two parts that are the bases of the invariant subspaces  $I_1$  and  $I_2$ . Then obviously the block  $A_3$  in (41) is also equal to zero and the matrix  $A$  has the quasi-diagonal form



of the matrix for a given operator in a given basis (Chapter III, p. 67), we find

$$L_1 = \begin{vmatrix} 0 & 0 & \dots & 0 & -\alpha_m \\ 1 & 0 & \dots & 0 & -\alpha_{m-1} \\ 0 & 1 & & & \\ \vdots & & \ddots & & \\ \vdots & & & 0 & -\alpha_2 \\ 0 & 0 & \dots & 1 & -\alpha_1 \end{vmatrix}. \quad (47)$$

Similarly

$$L_2 = \begin{vmatrix} 0 & 0 & \dots & 0 & -\beta_p \\ 1 & 0 & \dots & 0 & -\beta_{p-1} \\ 0 & 1 & & & \\ \vdots & & \ddots & & \\ \vdots & & & 0 & -\beta_1 \\ 0 & 0 & \dots & 1 & -\beta_1 \end{vmatrix}. \quad (48)$$

Computing the characteristic polynomials of the matrices  $L_1, L_2, \dots, L_t$ , we find:

$$|\lambda E - L_1| = \psi_1(\lambda), \quad |\lambda E - L_2| = \psi_2(\lambda), \quad \dots, \quad |\lambda E - L_t| = \psi_t(\lambda)$$

(for cyclic subspaces the characteristic polynomial of an operator  $A$  coincides with the minimal polynomial of the subspace relative to this operator).

The matrix  $L_1$  corresponds to the operator  $A$  in the 'canonical' basis (45). If  $A$  is the matrix corresponding to  $A$  in an arbitrary basis, then  $A$  is similar to  $L_1$ , i.e., there exists a non-singular matrix  $T$  such that

$$A = TL_1T^{-1}. \quad (49)$$

Of the matrix  $L_1$  we shall say that it has the *first natural normal form*. This form is characterized by

- 1) The quasi-diagonal form;
- 2) The special structure of the diagonal blocks (47), (48), etc.
- 3) The additional condition: the characteristic polynomial of each diagonal block is divisible by the characteristic polynomial of the following block.

If we start not from the second, but from the third decomposition theorem, then in exactly the same way we would obtain a matrix  $L_{II}$  corresponding to the operator  $A$  in the appropriate basis—a matrix having the *second natural normal form*, which is characterized by

- 1) The quasi-diagonal form

$$L_{II} = \{L^{(1)}, L^{(2)}, \dots, L^{(u)}\};$$

- 2) The special structure of the diagonal blocks (47), (48), etc.;
- 3) The additional condition: the characteristic polynomial of each block is a power of an irreducible polynomial over  $\mathbb{F}$ .

3. In the following section we shall show that in the class of similar matrices corresponding to one and the same operator there is one and only one matrix having the first normal form,<sup>9</sup> and one and only one<sup>10</sup> having the second normal form. Moreover, we shall give an algorithm for the computation of the polynomials  $\psi_1(\lambda), \psi_2(\lambda), \dots, \psi_l(\lambda)$  from the elements of the matrix  $A$ . Knowledge of these polynomials enables us to write out all the elements of the matrices  $L_I$  and  $L_{II}$  similar to  $A$  and having the first and second natural normal forms, respectively.

### § 6. Invariant Polynomials. Elementary Divisors

1. We<sup>11</sup> denote by  $D_p(\lambda)$  the greatest common divisor of all the minors of order  $p$  of the characteristic matrix  $A_\lambda = \lambda E - A$  ( $p = 1, 2, \dots, n$ ).<sup>12</sup> Since in the sequence

$$D_n(\lambda), D_{n-1}(\lambda), \dots, D_1(\lambda)$$

each polynomial is divisible by the following, the formulas

$$i_1(\lambda) = \frac{D_n(\lambda)}{D_{n-1}(\lambda)}, i_2(\lambda) = \frac{D_{n-1}(\lambda)}{D_{n-2}(\lambda)}, \dots, i_n(\lambda) = \frac{D_1(\lambda)}{D_0(\lambda)} \quad (D_0(\lambda) \equiv 1) \quad (50)$$

define  $n$  polynomials whose product is equal to the characteristic polynomial

$$\Delta(\lambda) = |\lambda E - A| = D_n(\lambda) = i_1(\lambda) i_2(\lambda) \dots i_n(\lambda). \quad (51)$$

We split the polynomials  $i_p(\lambda)$  ( $p = 1, 2, \dots, n$ ) into irreducible factors over  $\mathbb{F}$ :

$$i_p(\lambda) = [\varphi_1(\lambda)]^{r_p} [\varphi_2(\lambda)]^{s_p} \dots \quad (p = 1, 2, \dots, n); \quad (52)$$

where  $\varphi_1(\lambda), \varphi_2(\lambda), \dots$  are distinct irreducible polynomials over  $\mathbb{F}$ .

<sup>9</sup> This does not mean that there exists only one canonical basis of the form (45). There may be many canonical bases, but to all of them there corresponds one and the same matrix  $L_I$ .

<sup>10</sup> To within the order of the diagonal blocks.

<sup>11</sup> In subsection 1. of the present section we repeat the basic concepts of Chapter VI, § 3 for the characteristic matrix that were there established for an arbitrary polynomial matrix.

<sup>12</sup> We always take the highest coefficient of the greatest common divisor as 1.

The polynomials  $i_1(\lambda)$ ,  $i_2(\lambda)$ , ...,  $i_n(\lambda)$  are called the *invariant polynomials*, and all the non-constant powers among  $[\varphi_1(\lambda)]^{p_1}$ ,  $[\varphi_2(\lambda)]^{p_2}$ , ... are called the *elementary divisors*, of the characteristic matrix  $A_\lambda = \lambda E - A$  or, simply, of  $A$ .

The product of all the elementary divisors, like the product of all the invariant polynomials, is equal to the characteristic polynomial  $\Delta(\lambda) = |\lambda E - A|$ .

The name 'invariant polynomial' is justified by the fact that two similar matrices  $A$  and  $\tilde{A}$ ,

$$\tilde{A} = T^{-1}AT, \quad (53)$$

always have identical invariant polynomials

$$i_p(\lambda) = \tilde{i}_p(\lambda) \quad (p = 1, 2, \dots, n). \quad (54)$$

For it follows from (53) that

$$\tilde{A}_\lambda = \lambda E - \tilde{A} = T^{-1}(\lambda E - A)T = T^{-1}A_\lambda T. \quad (55)$$

Hence (see Chapter I, § 2) we obtain a relation between the minors of the similar matrices  $A_\lambda$  and  $\tilde{A}_\lambda$ :

$$\begin{aligned} \tilde{A}_\lambda \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \\ = \sum_{\substack{\alpha_1 < \alpha_2 < \dots < \alpha_p \\ \beta_1 < \beta_2 < \dots < \beta_p}} T^{-1} \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ \alpha_1 & \alpha_2 & \dots & \alpha_p \end{pmatrix} A_\lambda \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_p \\ \beta_1 & \beta_2 & \dots & \beta_p \end{pmatrix} T \begin{pmatrix} \beta_1 & \beta_2 & \dots & \beta_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \end{aligned} \quad (56)$$

$(p = 1, 2, \dots, n).$

This equation shows that every common divisor of all the minors of order  $p$  of  $A_\lambda$  is a common divisor of all the minors of order  $p$  of  $\tilde{A}_\lambda$ , and vice versa (since  $A$  and  $\tilde{A}$  can interchange places). Hence it follows that  $D_p(\lambda) = \tilde{D}_p(\lambda)$  ( $p = 1, 2, \dots, n$ ) and that (54) holds.

Since all the matrices representing a given operator  $A$  in various bases are similar and therefore have the same invariant polynomials and the same elementary divisors, we can speak of the invariant polynomials and the elementary divisors of an operator  $A$ .

**2.** We choose now for  $\tilde{A}$  the matrix  $L_1$  having the first natural normal form and we compute the invariant polynomials of  $A$  starting from the form of the matrix  $\tilde{A}_\lambda = \lambda E - \tilde{A}$  (in (57) this matrix is written out for the case  $m = 5$ ,  $p = 4$ ,  $q = 4$ ,  $r = 3$ ):

$$\left( \begin{array}{cccccccccccccccc}
 \lambda & 0 & 0 & 0 & \alpha_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \rightarrow 1 & \lambda & 0 & 0 & \alpha_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & -1 & \lambda & 0 & \alpha_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & -1 & \lambda & \alpha_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & -1 & \alpha_1 + \lambda & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \hline
 0 & 0 & 0 & 0 & 0 & \lambda & 0 & 0 & \beta_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & -1 & \lambda & 0 & \beta_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & -1 & \lambda & \beta_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & \beta_1 + \lambda & 0 & 0 & 0 & 0 & 0 & 0 \\
 \hline
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda & 0 & 0 & \gamma_4 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & \lambda & 0 & \gamma_3 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & \lambda & \gamma_2 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & \gamma_1 + \lambda & 0 & 0 & 0 \\
 \hline
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda & 0 & \varepsilon_3 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & \lambda & \varepsilon_2 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & \varepsilon_1 + \lambda
 \end{array} \right). \quad (57)$$

Using Laplace's Theorem, we find

$$D_n(\lambda) = |\lambda E - \tilde{A}| = |\lambda E - L_1| |\lambda E - L_2| \cdots |\lambda E - L_i| = \psi_1(\lambda) \psi_2(\lambda) \cdots \psi_i(\lambda). \quad (58)$$

Now let us find  $D_{n-1}(\lambda)$ . We consider the minor of the element  $\alpha_m$ . This minor, apart from a factor  $\pm 1$ , is equal to

$$|\lambda E - L_2| \cdots |\lambda E - L_i| = \psi_2(\lambda) \cdots \psi_i(\lambda). \quad (59)$$

We shall show that this minor of order  $n - 1$  is a divisor of all the other minors of order  $n - 1$ , so that

$$D_{n-1}(\lambda) = \psi_2(\lambda) \cdots \psi_i(\lambda). \quad (60)$$

For this purpose we first take the minor of an element outside the diagonal blocks and show that it vanishes. To obtain this minor we have to suppress one row and one column in the matrix (57). The lines crossed out in this case intersect two distinct diagonal blocks, so that in each of these blocks only one line is crossed out. Suppose, for example, that in the  $j$ -th diagonal block one of the rows is crossed out. In the minor we take that vertical strip which contains this diagonal block. In this strip, which has  $s$  columns, all the rows except  $s - 1$  rows consist entirely of zeros (we have denoted the order of  $A_j$  by  $s$ ). Expanding the determinant of order  $n - 1$  by Laplace's Theorem with respect to the minors of order  $s$  in this strip, we see that it is equal to zero.

Now we take the minor of an element inside one of the diagonal blocks. In this case the lines crossed out 'mutilate' only one of the diagonal blocks, say the  $j$ -th, and the matrix of the minor is again quasi-diagonal. Therefore the minor is equal to

$$\psi_1(\lambda) \cdots \psi_{j-1}(\lambda) \psi_{j+1}(\lambda) \cdots \psi_t(\lambda) \chi(\lambda), \quad (61)$$

where  $\chi(\lambda)$  is the determinant of the 'mutilated'  $j$ -th diagonal block. Since  $\psi_i(\lambda)$  is divisible by  $\psi_{i+1}(\lambda)$  ( $i=1, 2, \dots, t-1$ ), the product (61) is divisible by (59). Thus, equation (60) can be regarded as proved. By similar arguments we obtain:

$$\left. \begin{aligned} D_{n-2}(\lambda) &= \psi_3(\lambda) \cdots \psi_t(\lambda) \\ &\dots\dots\dots \\ D_{n-t+1}(\lambda) &= \psi_t(\lambda) \\ D_{n-t}(\lambda) &= \dots = D_1(\lambda) = 1 \end{aligned} \right\} \quad (62)$$

From (58), (60), and (62) we find:

$$\begin{aligned} \psi_1(\lambda) &= \frac{D_n(\lambda)}{D_{n-1}(\lambda)} = i_1(\lambda), \quad \psi_2(\lambda) = \frac{D_{n-1}(\lambda)}{D_{n-2}(\lambda)} = i_2(\lambda), \quad \dots \\ \psi_t(\lambda) &= \frac{D_{n-t+1}(\lambda)}{D_{n-t}(\lambda)} = i_t(\lambda), \\ i_{t+1}(\lambda) &= \dots = i_n(\lambda) = 1. \end{aligned} \quad (63)$$

The formulas (63) show that the polynomials  $\psi_1(\lambda)$ ,  $\psi_2(\lambda)$ ,  $\dots$ ,  $\psi_t(\lambda)$  coincide with the invariant polynomials, other than 1, of the operator  $A$  (or the corresponding matrix  $A$ ).

Let us give three equivalent formulations of the results obtained:

**THEOREM 9** (More precise form of the Second Decomposition Theorem): *If  $A$  is a linear operator in  $R$ , then the space  $R$  can be decomposed into cyclic subspaces*

$$R = I_1 + I_2 + \dots + I_t$$

*such that in the sequence of minimal polynomials  $\psi_1(\lambda)$ ,  $\psi_2(\lambda)$ ,  $\dots$ ,  $\psi_t(\lambda)$  of the subspaces  $I_1$ ,  $I_2$ ,  $\dots$ ,  $I_t$ , each is divisible by the following. The polynomials  $\psi_1(\lambda)$ ,  $\psi_2(\lambda)$ ,  $\dots$ ,  $\psi_t(\lambda)$  are uniquely determined: they coincide with the invariant polynomials, other than 1, of the operator  $A$ .*

**THEOREM 9'**: *For every linear operator  $A$  in  $R$  there exists a basis in which the matrix  $L_1$  that gives the operator is of the first natural normal form. This matrix is uniquely determined when the operator  $A$  is given: the characteristic polynomials of the diagonal blocks of  $L_1$  are the invariant polynomials of  $A$ .*

**THEOREM 9'':** *In every class of similar matrices (with elements in  $\mathbb{F}$ ) there exists one and only one matrix  $L_1$  having the first natural normal form. The characteristic polynomials of the diagonal blocks of  $L_1$  coincide with the invariant polynomials (other than 1) of every matrix of that class.*

On p. 194 we established that two similar matrices have the same invariant polynomials. Now suppose, conversely, that two matrices  $A$  and  $B$  with elements in  $\mathbb{F}$  are known to have the same invariant polynomials. Since the matrix  $L_1$  is uniquely determined when these polynomials are given, the two matrices  $A$  and  $B$  are similar to one and the same matrix  $L_1$  and, therefore, to each other. We thus arrive at the following proposition:

**THEOREM 10:** *Two matrices with elements in  $\mathbb{F}$  are similar if and only if they have the same invariant polynomials.<sup>13</sup>*

**3.** The characteristic polynomial  $\Delta(\lambda)$  of the operator  $A$  coincides with  $D_n(\lambda)$ , and hence with the product of all invariant polynomials:

$$\Delta(\lambda) = \psi_1(\lambda) \psi_2(\lambda) \cdots \psi_t(\lambda). \quad (64)$$

But  $\psi_1(\lambda)$  is the minimal polynomial of the whole space with respect to  $A$ ; hence  $\psi_1(A) = O$  and by (64)

$$\Delta(A) = O. \quad (65)$$

Thus we have incidentally obtained the Hamilton-Cayley Theorem (see Chapter IV, § 4):

*Every linear operator (every square matrix) satisfies its characteristic equation.*

In § 4 by splitting the polynomials  $\psi_1(\lambda), \psi_2(\lambda), \dots, \psi_t(\lambda)$  into irreducible factors over  $\mathbb{F}$ :

$$\begin{aligned} \psi_1(\lambda) &= [\varphi_1(\lambda)]^{c_1} [\varphi_2(\lambda)]^{c_2} \cdots [\varphi_s(\lambda)]^{c_s}, \\ \psi_2(\lambda) &= [\varphi_1(\lambda)]^{d_1} [\varphi_2(\lambda)]^{d_2} \cdots [\varphi_s(\lambda)]^{d_s}, \\ &\dots\dots\dots \\ \psi_t(\lambda) &= [\varphi_1(\lambda)]^{l_1} [\varphi_2(\lambda)]^{l_2} \cdots [\varphi_s(\lambda)]^{l_s} \end{aligned} \quad \left( \begin{array}{l} c_k \leq d_k \leq \dots \leq l_k, \\ k = 1, 2, \dots, s \end{array} \right), \quad (66)$$

we were led to the third decomposition theorem. To each power with non-zero exponent on the right-hand sides of (66) there corresponds an invariant subspace in this decomposition.

By (63) all the powers, other than 1, among  $[\varphi_k(\lambda)]^{c_k}, \dots, [\varphi_k(\lambda)]^{l_k}$  ( $k = 1, 2, \dots, s$ ) are the elementary divisors of  $A$  (or  $A$ ) in the field  $\mathbb{F}$  (see p. 194).

Thus we arrive at the following more precise statement of the third decomposition theorem:

<sup>13</sup> Or (what is the same) the same elementary divisors in the field  $\mathbb{F}$ .



**THEOREM 11:** *If  $A$  is a linear operator in a vector space  $R$  over a field  $F$ , then  $R$  can be split into cyclic subspaces whose minimal polynomials are the elementary divisors of  $A$  in  $F$ .*

Let

$$R = I' + I'' + \dots + I^{(u)} \quad (67)$$

be such a decomposition. We denote by  $e', e'', \dots, e^{(u)}$  generating vectors of the subspaces  $I', I'', \dots, I^{(u)}$  and from the 'cyclic' bases of these subspaces we form a basis of the whole space

$$e', Ae', \dots; e'', Ae'', \dots; e^{(u)}, Ae^{(u)}, \dots \quad (68)$$

It is easy to see that the matrix  $L_{II}$  corresponding to the operator  $A$  in the basis (68) has quasi-diagonal form, like  $L_I$ :

$$L_{II} = \{L_1, L_2, \dots, L_u\}. \quad (69)$$

The diagonal blocks  $L_1, L_2, \dots, L_u$  are of the same structure as the blocks (47) and (48) of  $L_I$ . However, the characteristic polynomials of these diagonal blocks are not the invariant polynomials, but the elementary divisors of  $A$ . The matrix  $L_{II}$  has the second natural normal form (see § 5).

We have arrived at another formulation of Theorem 11:

**THEOREM 11':** *For every linear operator  $A$  in  $R$  (over the field  $F$ ) there exists a basis in which the matrix  $L_{II}$  giving the operator is of the second natural normal form; the characteristic polynomials of the diagonal blocks are the elementary divisors of  $A$  in  $F$ .*

This theorem also admits a formulation in terms of matrices:

**THEOREM 11'':** *A matrix  $A$  with elements in the field  $F$  is always similar to a matrix  $L_{II}$  having the second natural normal form in which the characteristic polynomials of the diagonal blocks are the elementary divisors of  $A$ .*

Theorem 11 and the associated Theorems 11' and 11'' have, in a certain sense, a converse.

Let

$$R = I' + I'' + \dots + I^{(u)}$$

be an arbitrary decomposition of a space  $R$  into indecomposable invariant subspaces. Then by Theorem 7 the subspaces  $I', I'', \dots, I^{(u)}$  are cyclic and their minimal polynomials are powers of irreducible polynomials over  $F$ . We may write these powers, after adding powers with zero exponent if necessary, in the form<sup>14</sup>

<sup>14</sup> At least one of the numbers  $l_1, l_2, \dots, l_i$  is positive.

$$\begin{pmatrix} [\varphi_1(\lambda)]^{c_1}, [\varphi_2(\lambda)]^{c_2}, \dots, [\varphi_s(\lambda)]^{c_s}, \\ [\varphi_1(\lambda)]^{d_1}, [\varphi_2(\lambda)]^{d_2}, \dots, [\varphi_s(\lambda)]^{d_s}, \\ \dots \\ [\varphi_1(\lambda)]^{l_1}, [\varphi_2(\lambda)]^{l_2}, \dots, [\varphi_s(\lambda)]^{l_s} \end{pmatrix}, \quad \begin{pmatrix} c_k \geq d_k \geq \dots \geq l_k \geq 0, \\ k = 1, 2, \dots, s \end{pmatrix}. \quad (70)$$

We denote the sum of the subspaces whose minimal polynomials are in the first row by  $\mathbf{I}_1$ . Similarly, we introduce  $\mathbf{I}_2, \dots, \mathbf{I}_t$  ( $t$  is the number of rows in (70)). By Theorem 6, the subspaces  $\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_t$  are cyclic and their minimal polynomials  $\psi_1(\lambda), \psi_2(\lambda), \dots, \psi_t(\lambda)$  are determined by the formulas (66). Here in the sequence  $\psi_1(\lambda), \psi_2(\lambda), \dots, \psi_t(\lambda)$  each polynomial is divisible by the following. But then Theorem 9 is immediately applicable to the decomposition

$$R = I_1 + I_2 + \dots + I_t,$$

By this theorem

$$\psi_p(\lambda) = i_p(\lambda) \quad (p = 1, 2, \dots, n),$$

and therefore, by (66), all the powers (70) with non-zero exponent are the elementary divisors of  $A$  in the field  $F$ . Thus we have the following theorem:

**THEOREM 12:** *If the vector space  $\mathbf{R}$  (over the field  $\mathbf{F}$ ) is split in any way into decomposable invariant subspaces (with respect to an operator  $\mathbf{A}$ ), then the minimal polynomials of these subspaces are all the elementary divisors of  $\mathbf{A}$  in  $\mathbf{F}$ .*

There is an equivalent formulation in terms of matrices:

**THEOREM 12':** *In each class of similar matrices (with elements in  $\mathbf{F}$ ) there exists only one matrix (to within the order of the diagonal blocks) having the second normal form  $L_{II}$ ; the characteristic polynomials of its diagonal blocks are the elementary divisors of every matrix of the given class.*

Suppose that the space  $\mathbf{R}$  is split into two invariant subspaces (with respect to an operator  $\mathbf{A}$ )

$$\mathbf{R} = \mathbf{I}_1 + \mathbf{I}_2.$$

When we split  $I_1$  and  $I_2$  into indecomposable subspaces, we obtain at the same time a decomposition of the whole space  $R$  into indecomposable subspaces. Hence, bearing Theorem 12 in mind, we obtain:

**THEOREM 13:** *If the space  $\mathbf{R}$  is split into invariant subspaces with respect to an operator  $\mathbf{A}$ , then the elementary divisors of  $\mathbf{A}$  in each of these invariant subspaces, taken in their totality, form a complete system of elementary divisors of  $\mathbf{A}$  in  $\mathbf{R}$ .*

This theorem has the following matrix form:

**THEOREM 13':** *A complete system of elementary divisors in  $F$  of a quasi-diagonal matrix is obtained as the union of the elementary divisors of the diagonal blocks.*

Theorem 13' is often used for the actual process of finding the elementary divisors of a matrix.

## § 7. The Jordan Normal Form of a Matrix

**1.** Suppose that all the roots of the characteristic polynomial  $\Delta(\lambda)$  of an operator  $A$  belong to the field  $\mathbb{F}$ . This will hold true, in particular, if  $\mathbb{F}$  is the field of all complex numbers.

In this case, the decomposition of the invariant polynomials into elementary divisors in  $\mathbb{F}$  will look as follows:

$$\begin{aligned} i_1(\lambda) &= (\lambda - \lambda_1)^{\epsilon_1} (\lambda - \lambda_2)^{\epsilon_2} \dots (\lambda - \lambda_s)^{\epsilon_s}, \\ i_2(\lambda) &= (\lambda - \lambda_1)^{d_1} (\lambda - \lambda_2)^{d_2} \dots (\lambda - \lambda_s)^{d_s}, \\ &\vdots \\ i_s(\lambda) &= (\lambda - \lambda_1)^{l_1} (\lambda - \lambda_2)^{l_2} \dots (\lambda - \lambda_s)^{l_s}. \end{aligned} \quad \left( \begin{array}{l} c_k \geq d_k \geq \dots \geq l_k \geq 0, \\ c_k > 0; k = 1, 2, \dots, s \end{array} \right) \quad (71)$$

Since the product of all the invariant polynomials is equal to the characteristic polynomial  $\Delta(\lambda)$ ,  $\lambda_1, \lambda_2, \dots, \lambda_n$  in (7.1) are all the distinct roots of  $\Delta(\lambda)$ .

We take an arbitrary elementary divisor

$$(\lambda - \lambda_0)^p; \quad (72)$$

here  $\lambda_0$  is one of the numbers  $\lambda_1, \lambda_2, \dots, \lambda_s$  and  $p$  is one of the (non-zero) exponents  $c_k, d_k, \dots, l_k$  ( $k=1, 2, \dots, s$ ).

To this elementary divisor there corresponds in (67) a definite cyclic subspace  $I$ , generated by a vector which we denote by  $e$ . For this vector  $(\lambda - \lambda_0)^p$  is the minimal polynomial.

We consider the vectors

$$e_1 = (A - \lambda_0 E)^{p-1} e, \quad e_2 = (A - \lambda_0 E)^{p-2} e, \quad \dots, \quad e_p = e. \quad (73)$$

The vectors  $e_1, e_2, \dots, e_p$  are linearly independent, since otherwise there would be an annihilating polynomial for  $e$  of degree less than  $p$ , which is impossible. Now we note that

$$(A - \lambda_0 E) e_1 = 0, (A - \lambda_0 E) e_2 = e_1, \dots, (A - \lambda_0 E) e_p = e_{p-1} \quad (74)$$

**or**

$$Ae_1 = \lambda_0 e_1, Ae_2 = \lambda_0 e_2 + c_1, \dots, Ae_p = \lambda_0 e_p + e_{p-1}. \quad (75)$$

With the help of (75) we can easily write down the matrix corresponding to  $A$  in  $I$  for the basis (73). This matrix looks as follows:

$$\left\| \begin{array}{cccccc} \lambda_0 & 1 & 0 & \dots & 0 \\ 0 & \lambda_0 & 1 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & 1 \\ 0 & 0 & 0 & \dots & \lambda_0 \end{array} \right\| = \lambda_0 E^{(p)} + H^{(p)}, \quad (76)$$

where  $E^{(p)}$  is the unit matrix of order  $p$  and  $H^{(p)}$  the matrix of order  $p$  which has 1's along the first superdiagonal and 0's everywhere else.

Linearly independent vectors  $e_1, e_2, \dots, e_p$  for which (75) holds form a so-called *Jordan chain of vectors* in  $I$ . From Jordan chains connected with each subspace  $I', I'', \dots, I^{(u)}$  we form a *Jordan basis* of  $R$ . If we now denote the minimal polynomials of these subspaces, i.e., the elementary divisors of  $A$ , by

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots, (\lambda - \lambda_u)^{p_u} \quad (77)$$

(the numbers  $\lambda_1, \lambda_2, \dots, \lambda_u$  need not all be distinct), then the matrix  $J$  corresponding to  $A$  in a Jordan basis has the following quasi-diagonal form:

$$J = \{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \lambda_2 E^{(p_2)} + H^{(p_2)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)} \}. \quad (78)$$

We shall say of the matrix  $J$  that it is of *Jordan normal form* or simply *Jordan form*. The matrix  $J$  can be written down at once when the elementary divisors of  $A$  in the field  $F$  containing all the characteristic roots of the equation  $\Delta(\lambda) = 0$  are known.

*Every matrix  $A$  is similar to a matrix  $J$  of Jordan normal form*, i.e., for an arbitrary matrix  $A$  there always exists a non-singular matrix  $T$  ( $|T| \neq 0$ ) such that

$$A = TJT^{-1}.$$

If all the elementary divisors of  $A$  are of the first degree (and in that case only), the Jordan form is a diagonal matrix and we have:

$$A = T \{ \lambda_1, \lambda_2, \dots, \lambda_n \} T^{-1}.$$

Thus: *A linear operator  $A$  has simple structure* (see Chapter III, § 8) *if and only if all the elementary divisors of  $A$  are linear.*

Let us number the vectors  $e_1, e_2, \dots, e_p$  defined by (70) in the reverse order:

$$g_1 = e_p = e, g_2 = e_{p-1} = (A - \lambda_0 E) e, \dots, g_p = e_1 = (A - \lambda_0 E)^{p-1} e. \quad (79)$$

Then

$$(A - \lambda_0 E)g_1 = g_2, \quad (A - \lambda_0 E)g_2 = g_3, \quad \dots, \quad (A - \lambda_0 E)g_p = 0;$$

hence

$$Ag_1 = \lambda_0 g_1 + g_2, \quad Ag_2 = \lambda_0 g_2 + g_3, \quad \dots, \quad Ag_p = \lambda_0 g_p.$$

The vectors (79) form a basis in the cyclic invariant subspace  $I$  that corresponds in (67) to the elementary divisor  $(\lambda - \lambda_0)^p$ .

In this basis, as is easy to see, to the operator  $A$  there corresponds the matrix

$$\begin{vmatrix} \lambda_0 & 0 & 0 & \dots & 0 \\ 1 & \lambda_0 & 0 & \dots & 0 \\ 0 & 1 & \lambda_0 & & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & \lambda_0 \end{vmatrix} = \lambda_0 E^{(p)} + F^{(p)}.$$

We shall say of the vectors (79) that they form a *lower Jordan chain* of vectors. If we take a lower Jordan chain of vectors in each subspace  $I'$ ,  $I''$ ,  $\dots$ ,  $I^{(u)}$  of (67), we can form from these chains a *lower Jordan basis* in which to the operator  $A$  there corresponds the quasi-diagonal matrix

$$J_1 = \{ \lambda_1 E^{(p_1)} + F^{(p_1)}, \lambda_2 E^{(p_2)} + F^{(p_2)}, \dots, \lambda_u E^{(p_u)} + F^{(p_u)} \}. \quad (80)$$

We shall say of the matrix  $J_1$  that it is of *lower Jordan form*. In contrast to (80), we shall sometimes call (78) an *upper Jordan matrix*.

Thus: *Every matrix  $A$  is similar to an upper and to a lower Jordan matrix.*

## § 8. Krylov's Method of Transforming the Secular Equation

1. When a matrix  $A = \|a_{ik}\|_1^n$  is given, then its characteristic (secular) equation can be written in the form

$$\Delta(\lambda) \equiv (-1)^n \begin{vmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{vmatrix} = 0. \quad (81)$$

On the left-hand side of this equation is the characteristic polynomial  $\Delta(\lambda)$  of degree  $n$ . For the direct computation of the coefficients of this polynomial it is necessary to expand the characteristic determinant

$|A - \lambda E|$ ; and for large  $n$  this involves very cumbersome computational work, because  $\lambda$  occurs in the diagonal elements of the determinant.<sup>15</sup>

In 1937, A. N. Krylov [251] proposed a transformation of the characteristic determinant as a result of which  $\lambda$  occurs only in the elements of one column (or row).

Krylov's transformation simplifies the computation of the coefficients of the characteristic equation considerably.<sup>16</sup>

In this section we shall give an algebraic method of transforming the characteristic equation which differs somewhat from Krylov's own method.<sup>17</sup>

We consider an  $n$ -dimensional vector space  $\mathbf{R}$  with basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  and the linear operator  $A$  in  $\mathbf{R}$  determined by a given matrix  $A = \|a_{ik}\|_1^n$  in this basis. We take an arbitrary vector  $\mathbf{x} \neq \mathbf{o}$  in  $\mathbf{R}$  and form the sequence of vectors

$$\mathbf{x}, A\mathbf{x}, A^2\mathbf{x}, \dots \quad (82)$$

Suppose that the first  $p$  vectors  $\mathbf{x}, A\mathbf{x}, \dots, A^{p-1}\mathbf{x}$  of this sequence are linearly independent and that the  $(p+1)$ -st vector  $A^p\mathbf{x}$  is a linear combination of these  $p$  vectors:

$$A^p\mathbf{x} = -\alpha_p\mathbf{x} - \alpha_{p-1}A\mathbf{x} - \dots - \alpha_1A^{p-1}\mathbf{x} \quad (83)$$

or

$$\varphi(A)\mathbf{x} = \mathbf{o}, \quad (84)$$

where

$$\varphi(\lambda) = \lambda^p + \alpha_1\lambda^{p-1} + \dots + \alpha_p. \quad (85)$$

All the further vectors in (82) can also be expressed linearly by the first  $p$  vectors of the sequence.<sup>18</sup> Thus, in (82) there are  $p$  linearly independent

<sup>15</sup> We recall that the coefficient of  $\lambda^k$  in  $\mathcal{A}(\lambda)$  is equal (apart from the sign) to the sum of all the principal minors of order  $n-k$  in  $A$  ( $k=1, 2, \dots, n$ ). Thus, even for  $n=6$ , the direct determination of the coefficient of  $\lambda$  in  $\mathcal{A}(\lambda)$  would require the computation of six determinants of order 5; that of  $\lambda^2$  would require fifteen determinants of order 4; etc.

<sup>16</sup> The algebraic analysis of Krylov's method of transforming the secular equation is contained in a number of papers [268], [269], [211], [168], and [149].

<sup>17</sup> Krylov arrived at his method of transformation by starting from a system of  $n$  linear differential equations with constant coefficients. Krylov's approach in algebraic form can be found, for example, in [268] and [168] and in § 21 of the book [25].

<sup>18</sup> When we apply the operator  $A$  to both sides of (83) we express  $A^{p+1}\mathbf{x}$  linearly in terms of  $A\mathbf{x}, \dots, A^{p-1}\mathbf{x}, A^p\mathbf{x}$ . But  $A^p\mathbf{x}$ , by (83), is expressed linearly in terms of  $\mathbf{x}, A\mathbf{x}, \dots, A^{p-1}\mathbf{x}$ . Hence we obtain a similar expression for  $A^{p+1}\mathbf{x}$ . By applying the operator  $A$  to the expression thus obtained for  $A^{p+1}\mathbf{x}$ , we express  $A^{p+2}\mathbf{x}$  in terms of  $\mathbf{x}, A\mathbf{x}, \dots, A^{p+1}\mathbf{x}$ , etc.

vectors and this maximal number of linearly independent vectors in (82) is always realized by the *first*  $p$  vectors.

The polynomial  $q(\lambda)$  is the minimal (annihilating) polynomial of the vector  $\mathbf{x}$  with respect to the operator  $\mathbf{A}$  (see § 1). *The method of Krylov consists in an effective determination of the minimal polynomial  $q(\lambda)$  of  $\mathbf{x}$ .*

We consider separately two cases: the *regular* case, where  $p = n$ ; and the *singular* case, where  $p < n$ .

The polynomial  $q(\lambda)$  is a divisor of the minimal polynomial  $\psi(\lambda)$  of the whole space  $\mathbf{R}$ ,<sup>19</sup> and  $\psi(\lambda)$  in turn is a divisor of the characteristic polynomial  $\Delta(\lambda)$ . Therefore  $q(\lambda)$  is always a divisor of  $\Delta(\lambda)$ .

In the regular case,  $q(\lambda)$  and  $\Delta(\lambda)$  are of the same degree and, since their highest coefficients are equal, they coincide. Thus, in the regular case

$$\Delta(\lambda) \equiv \psi(\lambda) \equiv q(\lambda),$$

and therefore in the regular case Krylov's method is a method of computing the coefficients of the characteristic polynomial  $\Delta(\lambda)$ .

In the singular case, as we shall see later, Krylov's method does not enable us to determine  $\Delta(\lambda)$ , and in this case it only determines the divisor  $q(\lambda)$  of  $\Delta(\lambda)$ .

In explaining Krylov's transformation, we shall denote the coordinates of  $\mathbf{x}$  in the given basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  by  $a, b, \dots, l$ , and the coordinates of the vector  $\mathbf{A}^k \mathbf{x}$  by  $a_k, b_k, \dots, l_k$  ( $k = 1, 2, \dots, n$ ).

**2. Regular case:  $p = n$ .** In this case, the vectors  $\mathbf{x}, \mathbf{A}\mathbf{x}, \dots, \mathbf{A}^{n-1}\mathbf{x}$  are linearly independent and the equations (83), (84), and (85) assume the form

$$\mathbf{A}^n \mathbf{x} = -\alpha_n \mathbf{x} - \alpha_{n-1} \mathbf{A}\mathbf{x} - \dots - \alpha_1 \mathbf{A}^{n-1} \mathbf{x} \quad (86)$$

or

$$\Delta(\mathbf{A})\mathbf{x} = \mathbf{0}, \quad (87)$$

where

$$\Delta(\lambda) = \lambda^n + \alpha_1 \lambda^{n-1} + \dots + \alpha_{n-1} \lambda + \alpha_n. \quad (88)$$

The condition of linear independence of the vectors  $\mathbf{x}, \mathbf{A}\mathbf{x}, \dots, \mathbf{A}^{n-1}\mathbf{x}$  may be written analytically as follows (see Chapter III, § 1):

$$M = \begin{vmatrix} a & b & \dots & l \\ a_1 & b_1 & \dots & l_1 \\ \dots & \dots & \dots & \dots \\ a_{n-1} & b_{n-1} & \dots & l_{n-1} \end{vmatrix} \neq 0. \quad (89)$$

We consider the matrix formed from the coordinate vectors  $\mathbf{x}, \mathbf{A}\mathbf{x}, \dots, \mathbf{A}^{n-1}\mathbf{x}$ :

<sup>19</sup>  $\psi(\lambda)$  is the minimal polynomial of  $\mathbf{A}$ .





$$M\Delta(\lambda) = \begin{vmatrix} a & b & \dots & l & 1 \\ a_1 & b_1 & \dots & l_1 & \lambda \\ \dots & \dots & \dots & \dots & \dots \\ a_{n-1} & b_{n-1} & \dots & l_{n-1} & \lambda^{n-1} \\ a_n & b_n & \dots & l_n & \lambda^n \end{vmatrix}, \quad (93)$$

where the constant factor  $M$  is determined by (89) and differs from zero.

The identity (93) represents Krylov's transformation. In Krylov's determinant on the right-hand side of the identity,  $\lambda$  occurs only in the elements of the last column; the remaining elements of the determinant do not depend on  $\lambda$ .

*Note.* In the regular case, the whole space  $R$  is cyclic (with respect to  $A$ ). If we choose the vectors  $x, Ax, \dots, A^{n-1}x$  as a basis, then in this basis the operator  $A$  corresponds to a matrix  $\tilde{A}$  having the natural normal form

$$\tilde{A} = \begin{vmatrix} 0 & 0 & \dots & 0 & -\alpha_n \\ 1 & 0 & \dots & 0 & -\alpha_{n-1} \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & & \cdot \\ \cdot & \cdot & & 0 & -\alpha_2 \\ 0 & \dots & 1 & -\alpha_1 \end{vmatrix}. \quad (94)$$

The transition from the original basis  $e_1, e_2, \dots, e_n$  to the basis  $x, Ax, \dots, A^{n-1}x$  is accomplished by means of the non-singular transforming matrix

$$T = \begin{vmatrix} a & a_1 & \dots & a_{n-1} \\ b & b_1 & \dots & b_{n-1} \\ \dots & \dots & \dots & \dots \\ l & l_1 & \dots & l_{n-1} \end{vmatrix}. \quad (95)$$

and then

$$A = T\tilde{A}T^{-1}. \quad (96)$$

**3. Singular case:  $p < n$ .** In this case, the vectors  $x, Ax, \dots, A^{n-1}x$  are linearly dependent, so that

$$M = \begin{vmatrix} a & b & \dots & l \\ a_1 & b_1 & \dots & l_1 \\ \dots & \dots & \dots & \dots \\ a_{n-1} & b_{n-1} & \dots & l_{n-1} \end{vmatrix} = 0.$$



$$M^* \varphi(\lambda) = \begin{vmatrix} c & f & \dots & h & 1 \\ c_1 & f_1 & \dots & h_1 & \lambda \\ \dots & \dots & \dots & \dots & \dots \\ c_{p-1} & f_{p-1} & \dots & h_{p-1} & \lambda^{p-1} \\ c_p & f_p & \dots & h_p & \lambda^p \end{vmatrix}. \quad (100)$$

4. Let us now clarify the problem: for what matrices  $A = \|a_{ik}\|_1^n$  and for what choice of the original vector  $\mathbf{x}$  or, what is the same, of the initial parameters  $a, b, \dots, l$  the regular case holds.

We have seen that in the regular case

$$\Delta(\lambda) \equiv \psi(\lambda) \equiv \varphi(\lambda).$$

The fact that the characteristic polynomial  $\Delta(\lambda)$  coincides with the minimal polynomial  $\psi(\lambda)$  means that in the matrix  $A = \|a_{ik}\|_1^n$  there are no two elementary divisors with one and the same characteristic value, i.e., all the elementary divisors are co-prime in pairs. In the case where  $A$  is a matrix of simple structure, this requirement is equivalent to the condition that the characteristic equation of  $A$  have no multiple roots.

The fact that the polynomials  $\psi(\lambda)$  and  $\varphi(\lambda)$  coincide means that for  $\mathbf{x}$  we have chosen a vector that generates (by means of  $A$ ) the whole space  $\mathbf{R}$ . Such a vector always exists, by Theorem 2 of § 2.

But if the condition  $\Delta(\lambda) \equiv \psi(\lambda)$  is not satisfied, then however we choose the vector  $\mathbf{x} \neq \mathbf{o}$ , we do not obtain  $\Delta(\lambda)$ , since the polynomial  $\varphi(\lambda)$  obtained by Krylov's method is a divisor of  $\psi(\lambda)$  which in this case does not coincide with  $\Delta(\lambda)$  but is only a factor of it. By varying the vector  $\mathbf{x}$  we may obtain for  $\varphi(\lambda)$  every divisor of  $\psi(\lambda)$ .<sup>22</sup>

The results we have reached can be stated in the form of the following theorem:

**THEOREM 14:** *Krylov's transformation gives an expression for the characteristic polynomial  $\Delta(\lambda)$  of the matrix  $A = \|a_{ik}\|_1^n$  in the form of the determinant (93) if and only if two conditions are satisfied:*

1. *The elementary divisors of  $A$  are co-prime in pairs.*
2. *The initial parameters  $a, b, \dots, l$  are the coordinates of a vector  $\mathbf{x}$  that generates the whole  $n$ -dimensional space (by means of the operator  $A$  corresponding to the matrix  $A$ ).<sup>23</sup>*

<sup>22</sup> See, for example, [168], p. 48.

<sup>23</sup> In analytical form, this condition means that the columns  $\mathbf{x}, A\mathbf{x}, \dots, A^{n-1}\mathbf{x}$  are linearly independent, where  $\mathbf{x} = (a, b, \dots, l)$ .



Under the given matrix  $A$  we write the row of the coordinates of  $x$ :  $a, b, \dots, l$ . These numbers are given arbitrarily (with only one condition: at least one is different from zero). Under the row  $a, b, \dots, l$  we write the row  $a_1, b_1, \dots, l_1$ , i.e., the coordinates of the vector  $Ax$ . The numbers  $a_1, b_1, \dots, l_1$  are obtained by multiplying the row  $a, b, \dots, l$  successively into the rows of the given matrix  $A$ . For example,  $a_1 = a_{11}a + a_{12}b + \dots + a_{1n}l$ ,  $b_1 = a_{21}a + a_{22}b + \dots + a_{2n}l$ , etc. Under the row  $a_1, b_1, \dots, l_1$  we write the row  $a_2, b_2, \dots, l_2$ , etc. Each of the rows, beginning with the second, is determined by multiplying the preceding row successively into the rows of the given matrix.

Above the given matrix we write the sum row as a check.

$$\begin{array}{rcccl}
 & & 8 & 3 & -10 & -3 \\
 & & 3 & -1 & -4 & 2 \\
 A = & \left\| \begin{array}{ccccc} 2 & 3 & -2 & -4 \\ 2 & -1 & -3 & 2 \\ 1 & 2 & -1 & -3 \end{array} \right\| \\
 x = e_1 + e_2 & \left\{ \begin{array}{ccccc} 1 & 1 & 0 & 0 & -1 & 1 \\ Ax & 2 & 5 & 1 & 3 & -1 & -1 \\ A^2x & 3 & 5 & 2 & 2 & 1 & -1 \\ A^3x & 0 & 9 & -1 & 5 & 1 & 1 \\ A^4x & 5 & 9 & 4 & 4 & & \\ y & \left\{ \begin{array}{ccccc} 0 & 8 & 0 & 4 \\ 0 & 2 & 0 & 1 \\ -4 & 0 & -4 & 0 \\ 1 & 0 & 1 & 0 \end{array} \right.
 \end{array}
 \right.
 \end{array}$$

The given case is regular, because

$$M = \begin{vmatrix} 1 & 1 & 0 & 0 \\ 2 & 5 & 1 & 3 \\ 3 & 5 & 2 & 2 \\ 0 & 9 & -1 & 5 \end{vmatrix} = -16 \neq 0.$$

Krylov's determinant has the form

$$-16 \Delta(\lambda) = \begin{vmatrix} 1 & 1 & 0 & 0 & 1 \\ 2 & 5 & 1 & 3 & \lambda \\ 3 & 5 & 2 & 2 & \lambda^2 \\ 0 & 9 & -1 & 5 & \lambda^3 \\ 5 & 9 & 4 & 4 & \lambda^4 \end{vmatrix}.$$

Expanding this determinant and cancelling  $-16$  we find:

$$\Delta(\lambda) = \lambda^4 - 2\lambda^3 + 1 = (\lambda - 1)^2(\lambda + 1)^2.$$

We denote by

$$y = \xi_1 x + \xi_2 Ax + \xi_3 A^2 x + \xi_4 A^3 x$$

a characteristic vector of  $A$  corresponding to the characteristic value  $\lambda_0 = 1$ . We find the numbers  $\xi_1, \xi_2, \xi_3, \xi_4$  by the formulas (103):

$$\xi_4 = 1, \xi_3 = 1 \cdot \lambda_0 + 0 = 1, \xi_2 = 1 \cdot \lambda_0 - 2 = -1, \xi_1 = -1 \cdot \lambda_0 + 0 = -1.$$

The control equation  $-1 \cdot \lambda_0 + 1 = 0$  is, of course, satisfied.

We place the numbers  $\xi_1, \xi_2, \xi_3, \xi_4$  in a vertical column parallel to the columns of  $x, Ax, A^2x, A^3x$ . Multiplying the column  $\xi_1, \xi_2, \xi_3, \xi_4$  into the columns  $a_1, a_2, a_3, a_4$ , we obtain the first coordinate  $a'$  of the vector  $y$  in the original basis  $e_1, e_2, e_3, e_4$ ; similarly we obtain  $b', c', d'$ . As coordinates of  $y$  we find (after cancelling by 4): 0, 2, 0, 1. Similarly, we determine the coordinates 1, 0, 1, 0 of a characteristic vector  $z$  for the characteristic value  $\lambda_0 = -1$ .

Furthermore, by (94) and (95),

$$A = T \tilde{A} T^{-1}$$

where

$$\tilde{A} = \begin{vmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 0 \end{vmatrix}, \quad T = \begin{vmatrix} 1 & 2 & 3 & 0 \\ 1 & 5 & 5 & 9 \\ 0 & 1 & 2 & -1 \\ 0 & 3 & 2 & 5 \end{vmatrix}.$$

*Example 2.* We consider the same matrix  $A$ , but as initial parameters we take the numbers  $a = 1, b = 0, c = 0, d = 0$ .

$$A = \begin{vmatrix} 3 & 3 & -10 & -3 \\ 3 & -1 & -4 & 2 \\ 2 & 3 & -2 & -4 \\ 2 & -1 & -3 & 2 \\ 1 & 2 & -1 & -3 \end{vmatrix}$$

$$x = e_1 \quad \begin{vmatrix} 1 & 0 & 0 & 0 \\ Ax & 3 & 2 & 2 & 1 \\ A^2x & 1 & 4 & 0 & 2 \\ A^3x & 3 & 6 & 2 & 3 \end{vmatrix}$$

But in this case

$$M = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 3 & 2 & 2 & 1 \\ 1 & 4 & 0 & 2 \\ 3 & 6 & 2 & 3 \end{vmatrix} = 0$$

and  $p = 3$ . We have a singular case to deal with.

Taking the first three coordinates of the vectors  $\mathbf{x}$ ,  $A\mathbf{x}$ ,  $A^2\mathbf{x}$ ,  $A^3\mathbf{x}$ , we write the Krylov determinant in the form

$$\begin{vmatrix} 1 & 0 & 0 & 1 \\ 3 & 2 & 2 & \lambda \\ 1 & 4 & 0 & \lambda^2 \\ 3 & 6 & 2 & \lambda^3 \end{vmatrix}.$$

Expanding this determinant and cancelling  $-8$ , we obtain:

$$\varphi(\lambda) = \lambda^3 - \lambda^2 - \lambda + 1 = (\lambda - 1)^2(\lambda + 1).$$

Hence we find three characteristic values:  $\lambda_1 = 1$ ,  $\lambda_2 = 1$ ,  $\lambda_3 = -1$ . The fourth characteristic value can be obtained from the condition that the sum of all the characteristic values must be equal to the trace of the matrix. But  $\text{tr } A = 0$ . Hence  $\lambda_4 = -1$ .

These examples show that in applying Krylov's method, when we write down successively the rows of the matrix

$$\left\| \begin{array}{cccccc} a & b & \dots & l \\ a_1 & b_1 & \dots & l_1 \\ a_2 & b_2 & \dots & l_2 \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{array} \right\| \quad (105)$$

it is necessary to watch the rank of the matrix obtained so that we stop after the first row (the  $(p + 1)$ -st from above) that is a linear combination of the preceding ones. The determination of the rank is connected with the computation of certain determinants. Moreover, after obtaining Krylov's determinant in the form (93) or (100), in order to expand it with respect to the elements of the last column we have to compute a certain number of determinants of order  $p - 1$  (in the regular case, of order  $n - 1$ ).

Instead of expanding Krylov's determinant we can determine the coefficients  $a_1, a_2, \dots$  directly from the system of equations (91) (or (99)) by applying any efficient method of solution to the system—for example, the elimination method. This method can be applied immediately to the matrix

$$\left\| \begin{array}{cccccc} a & b & \dots & l & 1 \\ a_1 & b_1 & \dots & l_1 & \lambda \\ a_2 & b_2 & \dots & l_2 & \lambda^2 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \end{array} \right\| \quad (106)$$

by using it in parallel with the computation of the corresponding rows by Krylov's method. We shall then discover at once a row of the matrix (105)

that depends on the preceding ones, without computing any determinant.

Let us explain this in some detail. In the first row of (106) we take an arbitrary element  $c \neq 0$  and we use it to make the element  $c_1$  under it into zero, by subtracting from the second row the first row multiplied by  $c_1/c$ . Next we take an element  $f_1^* \neq 0$  in the second row and by means of  $c$  and  $f_1^*$  we make the elements  $c_2$  and  $f_2$  into zero, etc.<sup>25</sup> As a result of such a transformation, the element in the last column of (106) is replaced by a polynomial of degree  $k$ ,  $g_k(\lambda) = \lambda^k + \dots$  ( $k = 0, 1, 2, \dots$ )

Since under our transformation the rank of the matrix formed from the first  $k$  rows for any  $k$  and the first  $n$  columns of (106) does not change, the  $(p+1)$ -st row of the matrix must, after the transformation, have the form

$$0, 0, \dots, 0, g_p(\lambda).$$

Our transformation does not change the value of the Krylov determinant

$$\begin{vmatrix} c & f & \dots & h & 1 \\ c_1 & f_1 & \dots & h_1 & \lambda \\ \dots & \dots & \dots & \dots & \dots \\ c_{p-1} & f_{p-1} & \dots & h_{p-1} & \lambda^{p-1} \\ c_p & f_p & \dots & h_p & \lambda^p \end{vmatrix} = M^* \varphi(\lambda).$$

Therefore

$$M^* \varphi(\lambda) = c f_1^* \dots g_p(\lambda), \quad (107)$$

i.e.,<sup>26</sup>  $g_p(\lambda)$  is the required polynomial  $\varphi(\lambda)$ :  $g_p(\lambda) \equiv \varphi(\lambda)$ .

We recommend the following simplification. After obtaining the  $k$ -th transformed row of (106)

$$a_{k-1}^*, b_{k-1}^*, \dots, l_{k-1}^*, g_{k-1}(\lambda), \quad (108)$$

one should obtain the following  $(k+1)$ -st row by multiplying  $a_{k-1}^*, b_{k-1}^*, \dots, l_{k-1}^*$  (and not the original  $a_{k-1}, b_{k-1}, \dots, l_{k-1}$ ) into the rows of the given matrix.<sup>27</sup> Then we find the  $(k+1)$ -st row in the form

$$\tilde{a}_k, \tilde{b}_k, \dots, \tilde{l}_k, \lambda g_{k-1}(\lambda),$$

and after subtracting the preceding rows, we obtain:

<sup>25</sup> The elements  $c, f_1^*, \dots$  must not belong to the last column containing the powers of  $\lambda$ .

<sup>26</sup> We recall that the highest coefficients of  $\varphi(\lambda)$  and  $g_p(\lambda)$  are 1.

<sup>27</sup> The simplification consists in the fact that in the row of (108) to be transformed  $k-1$  elements are equal to zero. Therefore it is simple to multiply such a row into the rows of  $A$ .



$$a_k^*, b_k^*, \dots, l_k^*, g_k(\lambda).$$

The slight modification of Krylov's method that we have recommended (its combination with the elimination method) enables us to find at once the polynomial  $\varphi(\lambda)$  that we are interested in (in the regular case,  $\Delta(\lambda)$ ) without computing any determinants or solving any auxiliary system of equations.<sup>28</sup>

*Example.*

$$A = \begin{array}{ccccc|c} & 4 & 4 & 1 & 5 & 0 \\ \hline & 1 & 1 & -1 & 1 & 0 \\ & 1 & 2 & -1 & 0 & 1 \\ -1 & 2 & 3 & -1 & 0 & \\ \hline 1 & -2 & 1 & 2 & -1 & \\ 2 & 1 & -1 & 3 & 0 & \\ \hline 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & -1 & 0 & \lambda \\ 0 & 2 & 3 & -4 & -2 & \lambda^2 \quad [2 - 4\lambda] \\ 0 & -2 & 3 & 0 & 0 & \lambda^2 - 4\lambda + 2 \\ -5 & -7 & 5 & 7 & -5 & \lambda^3 - 4\lambda^2 + 2\lambda \quad [5 + 7\lambda] \\ -5 & 0 & 5 & 0 & 0 & \lambda^3 - 4\lambda^2 + 9\lambda + 5 \\ -10 & -10 & 20 & 0 & -15 & \lambda^4 - 4\lambda^3 + 9\lambda^2 + 5\lambda \quad [15 - 5(\lambda^2 - 4\lambda + 2) \\ & & & & & \quad - 2(\lambda^3 - 4\lambda^2 + 9\lambda + 5)] \\ 0 & 0 & -5 & 0 & 0 & \lambda^4 - 6\lambda^3 + 12\lambda^2 + 7\lambda - 5 \\ 5 & 5 & -15 & -5 & 5 & \lambda^5 - 6\lambda^4 + 12\lambda^3 + 7\lambda^2 - 5\lambda \quad [-5 - 5\lambda + (\lambda^2 \\ & & & & & \quad - 4\lambda^2 + 9\lambda + 5) - 2(\lambda^4 - 6\lambda^3 + 12\lambda^2 + 7\lambda - 5)] \\ 0 & 0 & 0 & 0 & 0 & \lambda^5 - 8\lambda^4 + 25\lambda^3 - 21\lambda^2 - 15\lambda + 10 \\ & & & & & \underbrace{\hspace{10em}}_{\Delta(\lambda)} \end{array}$$

<sup>28</sup> Apart from the method of Krylov, we have acquainted the reader in Chapter IV with the method of D. K. Faddeev for the computation of the coefficients of the characteristic polynomial. Faddeev's method involves more computations than Krylov's but it is more general, being without singular cases. We also refer the reader to the very effective method of A. M. Danilevskii [131]; see also the expository paper [376] and the book [15], § 24. See also [5] and [194].

## CHAPTER VIII

### MATRIX EQUATIONS

In this chapter we consider certain types of matrix equations that occur in various problems in the theory of matrices and its applications.

#### § 1. The Equation $AX = XB$

1. Suppose that the equation

$$AX = XB \quad (1)$$

is given, where  $A$  and  $B$  are square matrices (in general of different orders)

$$A = \|a_{ij}\|_1^m, \quad B = \|b_{kl}\|_1^n$$

and where  $X$  is an unknown rectangular matrix of dimension  $m \times n$ :

$$X = \|x_{jk}\| \quad (j=1, 2, \dots, m; \quad k=1, 2, \dots, n).$$

We write down the elementary divisors of  $A$  and  $B$  (in the field of complex numbers):

$$\begin{aligned} (A) : & (\lambda - \lambda_1)^{p_1}, \quad (\lambda - \lambda_2)^{p_2}, \quad \dots, \quad (\lambda - \lambda_u)^{p_u} \quad (p_1 + p_2 + \dots + p_u = m), \\ (B) : & (\lambda - \mu_1)^{q_1}, \quad (\lambda - \mu_2)^{q_2}, \quad \dots, \quad (\lambda - \mu_v)^{q_v} \quad (q_1 + q_2 + \dots + q_v = n). \end{aligned}$$

In accordance with these elementary divisors we reduce  $A$  and  $B$  to Jordan normal form

$$A = U\tilde{A}U^{-1}, \quad B = V\tilde{B}V^{-1}, \quad (2)$$

where  $U$  and  $V$  are square non-singular matrices of orders  $m$  and  $n$ , respectively, and  $\tilde{A}$  and  $\tilde{B}$  are the Jordan matrices:

$$\begin{aligned} \tilde{A} = & \{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \quad \lambda_2 E^{(p_2)} + H^{(p_2)}, \quad \dots, \quad \lambda_u E^{(p_u)} + H^{(p_u)} \}, \\ \tilde{B} = & \{ \mu_1 E^{(q_1)} + H^{(q_1)}, \quad \mu_2 E^{(q_2)} + H^{(q_2)}, \quad \dots, \quad \mu_v E^{(q_v)} + H^{(q_v)} \}. \end{aligned} \quad (3)$$

Replacing  $A$  and  $B$  in (1) by their expressions given in (2), we obtain:

$$U\tilde{A}U^{-1}X = XV\tilde{B}V^{-1}.$$

We multiply both sides of this equation on the left by  $U^{-1}$  and on the right by  $V$ :

$$\tilde{A}U^{-1}XV = U^{-1}XV\tilde{B}. \quad (4)$$

When we introduce in place of  $X$  a new unknown matrix  $\tilde{X}$  (of the same dimension  $m \times n$ )

$$\tilde{X} = U^{-1}XV, \quad (5)$$

we can write equation (4) as follows:

$$\tilde{A}\tilde{X} = \tilde{X}\tilde{B}. \quad (6)$$

We have thus replaced the matrix equation (1) by the equation (6), of the same form, in which the given matrices have Jordan normal form.

We partition  $\tilde{X}$  into blocks corresponding to the quasi-diagonal form of the matrices  $\tilde{A}$  and  $\tilde{B}$ :

$$\tilde{X} = (X_{\alpha\beta}) \quad (\alpha = 1, 2, \dots, u; \quad \beta = 1, 2, \dots, v)$$

(here  $X_{\alpha\beta}$  is a rectangular matrix of dimension  $p_\alpha \times q_\beta$  ( $\alpha = 1, 2, \dots, u$ ;  $\beta = 1, 2, \dots, v$ )).

Using the rule for multiplying a partitioned matrix by a quasi-diagonal one (see p. 42), we carry out the multiplication of the matrices on the left-hand and right-hand sides of (6). Then this equation breaks up into  $uv$  matrix equations

$$[\lambda_\alpha E^{(p_\alpha)} + H^{(p_\alpha)}] X_{\alpha\beta} = X_{\alpha\beta} [\mu_\beta E^{(q_\beta)} + H^{(q_\beta)}] \\ (\alpha = 1, 2, \dots, u; \quad \beta = 1, 2, \dots, v),$$

which we rewrite as follows:

$$(\mu_\beta - \lambda_\alpha) X_{\alpha\beta} = H_\alpha X_{\alpha\beta} - X_{\alpha\beta} G_\beta \quad (\alpha = 1, 2, \dots, u; \quad \beta = 1, 2, \dots, v); \quad (7)$$

we have used here the abbreviations

$$H_\alpha = H^{(p_\alpha)} \quad G_\beta = H^{(q_\beta)} \quad (\alpha = 1, 2, \dots, u; \quad \beta = 1, 2, \dots, v). \quad (8)$$

Let us take one of the equations (7). Two cases can occur:

1.  $\lambda_\alpha \neq \mu_\beta$ . We iterate equation (7)  $r-1$  times:<sup>1</sup>

<sup>1</sup> We multiply both sides of (7) by  $\mu_\beta - \lambda_\alpha$  and in each term of the right-hand side we replace  $(\mu_\beta - \lambda_\alpha) X_{\alpha\beta}$  by  $H_\alpha X_{\alpha\beta} - X_{\alpha\beta} G_\beta$ . This process is repeated  $r-1$  times

$$(\mu_\beta - \lambda_\alpha)^r X_{\alpha\beta} = \sum_{\sigma + \tau = r} (-1)^\tau \binom{r}{\tau} H_\alpha^\sigma X_{\alpha\beta} G_\beta^\tau. \quad (9)$$

Note that, by (8),

$$H_\alpha^{p_\alpha} = G_\beta^{q_\beta} = 0 \quad (10)$$

If in (9) we take  $r \geq p_\alpha + q_\beta - 1$ , then in each term of the sum on the right-hand side of (9) at least one of the relations

$$\sigma \geq p_\alpha, \quad \tau \geq q_\beta$$

is satisfied, so that by (10) either  $H_\alpha^\sigma = 0$  or  $G_\beta^\tau = 0$ . Moreover, since in this case  $\lambda_\alpha \neq \mu_\beta$ , we find from (9):

$$X_{\alpha\beta} = 0. \quad (11)$$

2.  $\lambda_\alpha = \mu_\beta$ . In this case equation (7) assumes the form

$$H_\alpha X_{\alpha\beta} = X_{\alpha\beta} G_\beta. \quad (12)$$

In the matrices  $H_\alpha$  and  $G_\beta$  the elements of the first superdiagonal are equal to 1, and all the remaining elements are zero. Taking this specific structure of  $H_\alpha$  and  $G_\beta$  into account and setting

$$X_{\alpha\beta} = \|\xi_{ik}\| \quad (i = 1, 2, \dots, p_\alpha; \quad k = 1, 2, \dots, q_\beta),$$

we replace the matrix equation (12) by the following equivalent system of scalar equations:<sup>2</sup>

$$\xi_{i+1,k} = \xi_{i,k-1} \quad (\xi_{i0} = \xi_{p_\alpha+1,k} = 0; \quad i = 1, 2, \dots, p_\alpha; \quad k = 1, 2, \dots, q_\beta). \quad (13)$$

The equations (13) have this meaning:

1) In the matrix  $X_{\alpha\beta}$  the elements of every line parallel to the main diagonal are equal;

$$2) \quad \xi_{21} = \xi_{31} = \dots = \xi_{p_\alpha 1} = \xi_{p_\alpha 2} = \dots = \xi_{p_\alpha, q_\beta-1} = 0.$$

Let  $p_\alpha = q_\beta$ . Then  $X_{\alpha\beta}$  is a square matrix. From 1) and 2) it follows that in  $X_{\alpha\beta}$  all the elements below the main diagonal are zero, all the elements in the main diagonal are equal to a certain number  $c_{\alpha\beta}$ , all the elements of the first superdiagonal are equal to a number  $c'_{\alpha\beta}$ , etc.; i.e.,

<sup>2</sup> From the structure of the matrices  $H_\alpha$  and  $G_\beta$  it follows that the product  $H_\alpha X_{\alpha\beta}$  is obtained from  $X_{\alpha\beta}$  by shifting all the rows one place upwards and filling the last row with zeros; similarly,  $X_{\alpha\beta} G_\beta$  is obtained from  $X_{\alpha\beta}$  by shifting all the columns one place to the right and filling the first column with zeros (see Chapter I, p. 14). To simplify the notation we do not write the additional indices  $\alpha, \beta$  in  $\xi_{ik}$ .

$$X_{\alpha\beta} = \left\| \begin{array}{cccc} c_{\alpha\beta} & c'_{\alpha\beta} & \dots & c_{\alpha\beta}^{(p_\alpha-1)} \\ 0 & c_{\alpha\beta} & & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & & \cdot & \cdot \\ 0 & \dots & 0 & c'_{\alpha\beta} \\ & & & c_{\alpha\beta} \end{array} \right\| = T_{p_\alpha}; \quad (14)$$

( $p_\alpha = q_\beta$ )

here  $c_{\alpha\beta}, c'_{\alpha\beta}, \dots, c_{\alpha\beta}^{(p_\alpha-1)}$  are arbitrary parameters (the equations (12) do not impose any restrictions on the values of these parameters).

It is easy to see that for  $p_\alpha < q_\beta$

$$X_{\alpha\beta} = \left( \begin{array}{c} \overbrace{0}^{q_\beta - p_\alpha} \\ 0 \end{array} \right), \quad T_{p_\alpha} \quad (15)$$

and for  $p_\alpha > q_\beta$

$$X_{\alpha\beta} = \left( \begin{array}{c} T_{q_\beta} \\ 0 \end{array} \right)_{p_\alpha - q_\beta}. \quad (16)$$

We shall say of the matrices (14), (15), and (16) that they have *regular* upper triangular form. The number of arbitrary parameters in  $X_{\alpha\beta}$  is equal to the smaller of the numbers  $p_\alpha$  and  $q_\beta$ . The scheme below shows the structure of the matrices  $X_{\alpha\beta}$  for  $x_\alpha = \mu_\beta$  (the arbitrary parameters are here denoted by  $a, b, c$ , and  $d$ ):

$$X_{\alpha\beta} = \left\| \begin{array}{cccc} a & b & c & d \\ 0 & a & b & c \\ 0 & 0 & a & b \\ 0 & 0 & 0 & a \end{array} \right\|, \quad X_{\alpha\beta} = \left\| \begin{array}{ccccc} 0 & 0 & a & b & c \\ 0 & 0 & 0 & a & b \\ 0 & 0 & 0 & 0 & a \end{array} \right\|, \quad X_{\alpha\beta} = \left\| \begin{array}{ccc} a & b & c \\ 0 & a & b \\ 0 & 0 & a \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right\|$$

( $p_\alpha = q_\beta = 4$ )                      ( $p_\alpha = 3, q_\beta = 5$ )                      ( $p_\alpha = 5, q_\beta = 3$ )

In order to subsume case 1 also in the count of arbitrary parameters in  $X$ , we denote by  $d_{\alpha\beta}(\lambda)$  the greatest common divisor of the elementary divisors  $(\lambda - \lambda_\alpha)^{p_\alpha}$  and  $(\lambda - \mu_\beta)^{q_\beta}$  and by  $\delta_{\alpha\beta}$  the degree of the polynomial  $d_{\alpha\beta}(\lambda)$  ( $\alpha = 1, 2, \dots, u$ ;  $\beta = 1, 2, \dots, v$ ). In case 1, we have  $\delta_{\alpha\beta} = 0$ ; in case 2,  $\delta_{\alpha\beta} = \min(p_\alpha, q_\beta)$ . Thus, in both cases the number of arbitrary parameters in  $X_{\alpha\beta}$  is equal to  $\delta_{\alpha\beta}$ . The number of arbitrary parameters in  $\tilde{X}$  is determined by the formula.

$$N = \sum_{\alpha=1}^u \sum_{\beta=1}^v \delta_{\alpha\beta}.$$

In what follows it will be convenient to denote the general solution of (6) by  $X_{\tilde{A}\tilde{B}}$  (so far we have denoted it by  $\tilde{X}$ ).

The results obtained in this section can be stated in the form of the following theorem:

THEOREM 1: *The general solution of the matrix equation*

$$AX = XB$$

where

$$A = \|a_{ik}\|_1^m = U\tilde{A}U^{-1} = U\{\lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)}\}U^{-1},$$

$$B = \|b_{ik}\|_1^n = V\tilde{B}V^{-1} = V\{\mu_1 E^{(q_1)} + H^{(q_1)}, \dots, \mu_v E^{(q_v)} + H^{(q_v)}\}V^{-1}$$

is given by the formula

$$X = UX_{\tilde{A}\tilde{B}}V^{-1} \quad (17)$$

Here  $X_{\tilde{A}\tilde{B}}$  is the general solution of the equation

$$\tilde{A}\tilde{X} = \tilde{X}\tilde{B}$$

and has the following structure:

$X_{\tilde{A}\tilde{B}}$  is decomposed into blocks

$$X_{\tilde{A}\tilde{B}} = (\tilde{X}_{\alpha\beta})_{p_\alpha}^{q_\beta} \quad (\alpha = 1, 2, \dots, u; \beta = 1, 2, \dots, v);$$

if  $\lambda_\alpha \neq \mu_\beta$ , then the null matrix stands in the place  $X_{\alpha\beta}$ , but if  $\lambda_\alpha = \mu_\beta$ , then an arbitrary regular upper triangular matrix stands in the place  $X_{\alpha\beta}$ .

$X_{\tilde{A}\tilde{B}}$ , and therefore also  $X$ , depends linearly on  $N$  arbitrary parameters  $c_1, c_2, \dots, c_N$

$$X = \sum_{j=1}^N c_j X_j, \quad (18)$$

where  $N$  is determined by the formula

$$N = \sum_{\alpha=1}^u \sum_{\beta=1}^v \delta_{\alpha\beta} \quad (19)$$

(here  $\delta_{\alpha\beta}$  denotes the degree of the greatest common divisor of  $(\lambda - \lambda_\alpha)^{p_\alpha}$  and  $(\lambda - \mu_\beta)^{q_\beta}$ ).

Note that the matrices  $X_1, X_2, \dots, X_N$  that occur in (18) are solutions of the original equation (1) ( $X_j$  is obtained from  $X$  by giving to the parameter  $c_j$  the value 1 and to the remaining parameters the value 0;  $j = 1, 2, \dots, N$ ). These solutions are linearly independent, since otherwise for certain values of the parameters  $c_1, c_2, \dots, c_N$ , not all zero, the matrix  $X$ , and therefore  $X_{\tilde{A}\tilde{B}}$ , would be the null matrix, which is impossible. Thus (18) shows that every solution of the original equation is a linear combination of  $N$  linearly independent solutions.

If the matrices  $A$  and  $B$  do not have common characteristic values (if the characteristic polynomials  $|\lambda E - A|$  and  $|\lambda E - B|$  are co-prime), then  $N = \sum_{\alpha=1}^u \sum_{\beta=1}^v \delta_{\alpha\beta} = 0$ , and so  $X = O$ , i.e., in this case the equation (1) has only the trivial solution  $X = O$ .

*Note.* Suppose that the elements of  $A$  and  $B$  belong to some number field  $\mathbb{F}$ . Then we cannot say that the elements of  $U$ ,  $V$ , and  $X_{\tilde{A}\tilde{B}}$  that occur in (17) also belong to  $\mathbb{F}$ . The elements of these matrices may be taken in an extension field  $\mathbb{F}_1$  which is obtained from  $\mathbb{F}$  by adjoining the roots of the characteristic equations  $|\lambda E - A| = 0$  and  $|\lambda E - B| = 0$ . We always have to deal with such an extension of the ground field when we use the reduction of given matrices to Jordan normal form.

However, the matrix equation (1) is equivalent to a system of  $mn$  linear homogeneous equations, where the unknown are the elements  $x_{jk}$  ( $j = 1, 2, 3, \dots, m; k = 1, 2, \dots, n$ ) of the required matrix  $X$ :

$$\sum_{j=1}^m a_{ij} x_{jk} = \sum_{h=1}^n x_{ih} b_{hk} \quad (i = 1, 2, \dots, m; k = 1, 2, \dots, n). \quad (20)$$

What we have shown is that this system has  $N$  linearly independent solutions, where  $N$  is determined by (19). But it is well known that fundamental linearly independent solutions can be chosen in the ground field  $\mathbb{F}$  to which the coefficients of (20) belong. Thus, in (18) the matrices  $X_1, X_2, \dots, X_N$  can be so chosen that their elements lie in  $\mathbb{F}$ . If we then give to the arbitrary parameters in (18) all possible values in  $\mathbb{F}$ , we obtain all the matrices  $X$  with elements in  $\mathbb{F}$  that satisfy the equation (1).<sup>3</sup>

## § 2. The Special Case $A = B$ . Commuting Matrices

1. Let us consider the special case of the equation (1):

$$AX = XA, \quad (21)$$

where  $A = \|a_{ik}\|_1^n$  is a given matrix and  $X = \|x_{ik}\|_1^n$  an unknown matrix. We have come to a problem of Frobenius: to determine all the matrices  $X$  that commute with a given matrix  $A$ .

We reduce  $A$  to Jordan normal form:

$$A = U \tilde{A} U^{-1} = U \{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)} \} U^{-1}. \quad (22)$$

<sup>3</sup> The matrices  $A = \|a_{ij}\|_1^m$  and  $B = \|b_{kl}\|_1^n$  determine a linear operator  $\hat{F}(X) = AX - XB$  in the space of rectangular matrices  $X$  of dimension  $m \times n$ . A treatment of operators of this type is contained in the paper [179].

Then when we set in (17)  $V = U$ ,  $\tilde{B} = \tilde{A}$  and denote  $X_{\tilde{A}\tilde{A}}$  simply by  $X_{\tilde{A}}$ , we obtain all solutions of (21), i.e., all matrices that commute with  $A$ , in the following form:

$$X = UX_{\tilde{A}}U^{-1}, \quad (23)$$

where  $X_{\tilde{A}}$  denotes an arbitrary matrix permutable with  $\tilde{A}$ . As we have explained in the preceding section,  $X_{\tilde{A}}$  is split into  $u^2$  blocks

$$X_{\tilde{A}} = (X_{\alpha\beta})_{\alpha, \beta=1}^u$$

corresponding to the splitting of the Jordan matrix  $\tilde{A}$  into blocks;  $X_{\alpha\beta}$  is either the null matrix or an arbitrary regular upper triangular matrix, depending on whether  $\lambda_\alpha \neq \lambda_\beta$  or  $\lambda_\alpha = \lambda_\beta$ .

As an example, we write down the elements of  $X_{\tilde{A}}$  in the case where  $A$  has the following elementary divisors:

$$(\lambda - \lambda_1)^4, (\lambda - \lambda_1)^3, (\lambda - \lambda_2)^2, \lambda - \lambda_2 \quad (\lambda_1 \neq \lambda_2).$$

In this case  $X_{\tilde{A}}$  has the following form:

$$\left\| \begin{array}{cccccccccccc} a & b & c & d & e & f & g & 0 & 0 & 0 & 0 \\ 0 & a & b & c & 0 & e & f & 0 & 0 & 0 & 0 \\ 0 & 0 & a & b & 0 & 0 & e & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & h & k & l & m & p & q & 0 & 0 & 0 & 0 \\ 0 & 0 & h & k & 0 & m & p & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & h & 0 & 0 & m & 0 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & r & s & t & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & r & 0 & \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w & z & \end{array} \right\| \quad (a, b, \dots, z \text{ are arbitrary parameters}).$$

The number of parameters in  $X_{\tilde{A}}$  is equal to  $N$ , where  $N = \sum_{\alpha, \beta=1}^u \delta_{\alpha\beta}$ ; here  $\delta_{\alpha\beta}$  denotes the degree of the greatest common divisor of the polynomials  $(\lambda - \lambda_\alpha)^{p_\alpha}$  and  $(\lambda - \lambda_\beta)^{p_\beta}$ .

Let us bring the invariant polynomials of  $A$  into the discussion:  $i_1(\lambda), i_2(\lambda), \dots, i_t(\lambda); i_{t+1}(\lambda) = \dots = i_n(\lambda) = 1$ . We denote the degrees of these polynomials by  $n_1 \geq n_2 \geq \dots \geq n_t > n_{t+1} = \dots = 0$ . Since each invariant polynomial is a product of certain co-prime elementary divisors, the formula for  $N$  can be written as follows:



$$N = \sum_{g,j=1}^t \kappa_{gj}, \quad (24)$$

where  $\kappa_{gj}$  is the degree of the greatest common divisor of  $i_g(\lambda)$  and  $i_j(\lambda)$  ( $g, j = 1, 2, \dots, t$ ). But the greatest common divisor of  $i_g(\lambda)$  and  $i_j(\lambda)$  is one of these polynomials and therefore  $\kappa_{gj} = \min(n_g, n_j)$ . Hence we obtain:

$$N = n_1 + 3n_2 + \dots + (2t-1)n_t.$$

$N$  is the number of linearly independent matrices that commute with  $A$  (we may assume that the elements of these matrices belong to the ground field  $\mathbb{F}$  containing the elements of  $A$ ; see the remark at the end of the preceding section). We have arrived at the following theorem:

**THEOREM 2:** *The number of linearly independent matrices that commute with the matrix  $A = \|a_{ik}\|_1^n$  is given by the formula*

$$N = n_1 + 3n_2 + \dots + (2t-1)n_t. \quad (25)$$

where  $n_1, n_2, \dots, n_t$  are the degrees of the non-constant invariant polynomials  $i_1(\lambda), i_2(\lambda), \dots, i_t(\lambda)$  of  $A$ .

Note that

$$n = n_1 + n_2 + \dots + n_t. \quad (26)$$

From (25) and (26) it follows that

$$N \geq n, \quad (27)$$

where the equality sign holds if and only if  $t=1$ , i.e., if all the elementary divisors of  $A$  are co-prime in pairs.

**2.** Let  $g(\lambda)$  be an arbitrary polynomial in  $\lambda$ . Then  $g(A)$  is permutable with  $A$ . There arises the converse question: when can every matrix that is permutable with  $A$  be expressed as a polynomial in  $A$ ? Every matrix that commutes with  $A$  would then be a linear combination of the linearly independent matrices

$$E, A, A^2, \dots, A^{n_1-1}.$$

Hence  $N = n_1 \leq n$ ; on comparing this with (27), we obtain:  $N = n_1 = n$ .

**COROLLARY 1 TO THEOREM 2:** *All the matrices that are permutable with  $A$  can be expressed as polynomials in  $A$  if and only if  $n_1 = n$ , i.e., if all the elementary divisors of  $A$  are co-prime in pairs.*

3. The polynomials in a matrix that commutes with  $A$  also commute with  $A$ . We raise the question: when can all the matrices that commute with  $A$  be expressed in the form of polynomials in one and the same matrix  $C$ ? Let us consider the case in which they can be so expressed. Then since by the Hamilton Cayley Theorem the matrix  $C$  satisfies its characteristic equation, every matrix that commutes with  $C$  must be expressible linearly by the matrices

$$E, C, C^2, \dots, C^{n-1}.$$

Therefore in this case  $N \leq n$ . Comparing this with (27), we find that  $N = n$ . Hence from (25) and (26) we also have  $n_1 = n$ .

COROLLARY 2 TO THEOREM 2: *All the matrices that are permutable with  $A$  can be expressed in the form of polynomials in one and the same matrix  $C$  if and only if  $n_1 = n$ , i.e. if and only if all the elementary divisors of  $\lambda E - A$  are co-prime. In this case all the matrices that are permutable with  $A$  can be represented in the form of polynomials in  $A$ .*

4. We mention a very important property of permutable matrices.

THEOREM 3: *If two matrices  $A = \| a_{ik} \|_1^n$  and  $B = \| b_{ik} \|_1^n$  are permutable and if one of them, say  $A$ , has quasi-diagonal form*

$$A = \{\overset{s_1}{\hat{A}_1}, \overset{s_2}{\hat{A}_2}\}, \quad (28)$$

where the matrices  $A_1$  and  $A_2$  do not have characteristic values in common, then the other matrix also has the same quasi-diagonal form

$$B = \{\overset{s_1}{\hat{B}_1}, \overset{s_2}{\hat{B}_2}\}. \quad (29)$$

*Proof.* We split  $B$  into blocks corresponding to the quasi-diagonal form (28):

$$B = \begin{pmatrix} \overset{s_1}{\hat{B}_1} & \overset{s_2}{\hat{X}} \\ Y & B_2 \end{pmatrix}.$$

From the relation  $AB = BA$  we obtain four matrix equations:

$$1. A_1 B_1 = B_1 A_1, \quad 2. A_1 X = X A_2, \quad 3. A_2 Y = Y A_1, \quad 4. A_2 B_2 = B_2 A_2. \quad (30)$$

As we explained in § 1 (p. 220), the second and third of the equations in (30) only have the solutions  $X = O$ ,  $Y = O$ , since  $A_1$  and  $A_2$  have no characteristic values in common. This proves our statement. The first and fourth of the equations in (30) express the permutability of  $A_1$  and  $B_1$  and of  $A_2$  and  $B_2$ .

In geometrical language, this theorem runs as follows:

**THEOREM 3':** *If*

$$\mathbf{R} = \mathbf{I}_1 + \mathbf{I}_2$$

*is a decomposition of the whole space  $\mathbf{R}$  into invariant subspaces  $\mathbf{I}_1$  and  $\mathbf{I}_2$  with respect to an operator  $\mathbf{A}$  and if the minimal polynomials of these subspaces (with respect to  $\mathbf{A}$ ) are co-prime, then  $\mathbf{I}_1$  and  $\mathbf{I}_2$  are invariant with respect to any linear operator  $\mathbf{B}$  that commutes with  $\mathbf{A}$ .*

Let us also give a geometrical proof of this statement. We denote by  $\psi_1(\lambda)$  and  $\psi_2(\lambda)$  the minimal polynomials of  $\mathbf{I}_1$  and  $\mathbf{I}_2$  with respect to  $\mathbf{A}$ . From the fact that they are co-prime it follows that all the vectors of  $\mathbf{R}$  that satisfy the equation  $\psi_1(\mathbf{A})\mathbf{x} = \mathbf{o}$  belong to  $\mathbf{I}_1$  and all the vectors that satisfy  $\psi_2(\mathbf{A})\mathbf{x} = \mathbf{o}$  belong to  $\mathbf{I}_2$ .<sup>4</sup> Let  $\mathbf{x}_1 \in \mathbf{I}_1$ . Then  $\psi_1(\mathbf{A})\mathbf{x}_1 = \mathbf{o}$ . The permutability of  $\mathbf{A}$  and  $\mathbf{B}$  implies that of  $\psi_1(\mathbf{A})$  and  $\mathbf{B}$ , so that

$$\psi_1(\mathbf{A})\mathbf{B}\mathbf{x}_1 = \mathbf{B}\psi_1(\mathbf{A})\mathbf{x}_1 = \mathbf{o},$$

i.e.,  $\mathbf{B}\mathbf{x}_1 \in \mathbf{I}_1$ . The invariance of  $\mathbf{I}_2$  with respect to  $\mathbf{B}$  is proved similarly.

This theorem leads to a number of corollaries:

**COROLLARY 1:** *If the linear operators  $\mathbf{A}, \mathbf{B}, \dots, \mathbf{L}$  are pairwise permutable, then the whole space  $\mathbf{R}$  can be split into subspaces invariant with respect to all the operators  $\mathbf{A}, \mathbf{B}, \dots, \mathbf{L}$*

$$\mathbf{R} = \mathbf{I}_1 + \mathbf{I}_2 + \dots + \mathbf{I}_w$$

*such that the minimal polynomial of each of these subspaces with respect to any one of the operators  $\mathbf{A}, \mathbf{B}, \dots, \mathbf{L}$  is a power of an irreducible polynomial.*

As a special case of this we obtain:

**COROLLARY 2:** *If the linear operators  $\mathbf{A}, \mathbf{B}, \dots, \mathbf{L}$  are pairwise permutable and all the characteristic values of these operators belong to the ground field, then the whole space  $\mathbf{R}$  can be split into subspaces  $\mathbf{I}_1, \mathbf{I}_2, \dots, \mathbf{I}_w$ , invariant with respect to all the operators such that each operator  $\mathbf{A}, \mathbf{B}, \dots, \mathbf{L}$  has equal characteristic values in each of them.*

Finally, we mention a further special case of this statement:

**COROLLARY 3:** *If  $\mathbf{A}, \mathbf{B}, \dots, \mathbf{L}$  are pairwise permutable operators of simple structure (see Chapter III, § 8), then a basis of the space can be formed from common characteristic vectors of these operators.*

We also give the matrix form of the last statement:

*Permutable matrices of simple structure can be brought into diagonal form simultaneously by a similarity transformation.*

<sup>4</sup> See Theorem 1 of Chapter VII (p. 179).

§ 3. The Equation  $AX - XB = C$ 

1. Suppose that the matrix equation

$$AX - XB = C \quad (31)$$

is given, where  $A = \| a_{ij} \|_1^m$  and  $B = \| b_{kl} \|_1^n$  are given square matrices of order  $m$  and  $n$  and where  $C = \| c_{ik} \|$  and  $X = \| x_{jk} \|$  are a given and an unknown rectangular matrix, respectively, of dimension  $m \times n$ . The equation (31) is equivalent to a system of  $mn$  scalar equations in the elements of  $X$ :

$$\sum_{j=1}^m a_{ij} x_{jk} - \sum_{l=1}^n x_{il} b_{lk} = c_{ik} \quad (i = 1, 2, \dots, m; k = 1, 2, \dots, n). \quad (31')$$

The corresponding homogeneous system of equations

$$\sum_{j=1}^m a_{ij} x_{jk} - \sum_{l=1}^n x_{il} b_{lk} = 0 \quad (i = 1, 2, \dots, m; k = 1, 2, \dots, n),$$

can be written in matrix form as follows:

$$AX - XB = 0. \quad (32)$$

Thus, if (32) only has the trivial solution  $X = 0$ , then (31) has a unique solution. But we have established in § 1 that the only solution of (32) is the trivial one if and only if  $A$  and  $B$  do not have common characteristic values. Therefore, *if the matrices  $A$  and  $B$  do not have characteristic values in common, then (31) has a unique solution; but if the matrices  $A$  and  $B$  have characteristic values in common, then two cases may arise depending on the 'constant' term  $C$ : either the equation (31) is contradictory, or it has an infinite number of solutions given by the formula*

$$X = X_0 + X_1,$$

where  $X_0$  is a fixed particular solution of (31) and  $X_1$  the general solution of the homogeneous equation (32) (the structure of  $X_1$  was described in § 1).

§ 4. The Scalar Equation  $f(X) = 0$ 

1. To begin with, let us consider the equation

$$g(X) = 0, \quad (33)$$

where

$$g(\lambda) = (\lambda - \lambda_1)^{\alpha_1} (\lambda - \lambda_2)^{\alpha_2} \dots (\lambda - \lambda_h)^{\alpha_h}$$

is a given polynomial in the variable  $\lambda$  and  $X$  is an unknown square matrix of order  $n$ . Since the minimal polynomial of  $X$ , i.e., the first invariant polynomial  $i_1(\lambda)$ , must be a divisor of  $g(\lambda)$ , the elementary divisors of  $X$  must have the following form:

$$(\lambda - \lambda_{j_1})^{p_{j_1}}, (\lambda - \lambda_{j_2})^{p_{j_2}}, \dots, (\lambda - \lambda_{j_\nu})^{p_{j_\nu}} \left( \begin{array}{l} j_1, j_2, \dots, j_\nu = 1, 2, \dots, h, \\ p_{j_1} \leq a_{j_1}, p_{j_2} \leq a_{j_2}, \dots, p_{j_\nu} \leq a_{j_\nu}, \\ p_{j_1} + p_{j_2} + \dots + p_{j_\nu} = n \end{array} \right)$$

(among the indices  $j_1, j_2, \dots, j_\nu$  there may be some that are equal;  $n$  is the given order of the unknown matrix  $X$ ).

We represent  $X$  in the form

$$X = T \{ \lambda_{j_1} E^{(p_{j_1})} + H^{(p_{j_1})}, \dots, \lambda_{j_\nu} E^{(p_{j_\nu})} + H^{(p_{j_\nu})} \} T^{-1}, \quad (34)$$

where  $T$  is an arbitrary non-singular matrix of order  $n$ . The set of solutions of the equation (33) with a given order of the unknown matrix splits, by formula (34), into a finite number of classes of similar matrices.

*Example 1.* Let the equation

$$X^m = O \quad (35)$$

be given.

If a certain power of a matrix is the null matrix, then the matrix is called *nilpotent*. The least exponent for which the power of the matrix is the null matrix is called the *index of nilpotency*.

Obviously, the solutions of (35) are all the nilpotent matrices with an index of nilpotency  $\mu \leq m$ . The formula that comprises all the solutions of a given order  $n$  looks as follows ( $T$  is an arbitrary non-singular matrix):

$$X = T \{ H^{(p_1)}, H^{(p_2)}, \dots, H^{(p_\nu)} \} T^{-1} \left( \begin{array}{l} p_1, p_2, \dots, p_\nu \leq m, \\ p_1 + p_2 + \dots + p_\nu = n \end{array} \right). \quad (36)$$

*Example 2.* Let the equation

$$X^2 = X \quad (37)$$

be given.

A matrix satisfying this equation is called *idempotent*. The elementary divisors of an idempotent matrix can only be  $\lambda$  or  $\lambda - 1$ . Therefore an idempotent matrix can be described as a matrix of simple structure (i.e., reducible to diagonal form) with characteristic values 0 or 1. The formula comprising all the idempotent matrices of a given order  $n$  has the form

$$X = T \underbrace{\{ 1, 1, \dots, 1, 0, \dots, 0 \}}_n T^{-1}, \quad (38)$$

where  $T$  is an arbitrary non-singular matrix of order  $n$ .

2. Let us now consider the more general equation

$$f(X) = O, \quad (39)$$

where  $f(\lambda)$  is a regular function of  $\lambda$  in some domain  $G$  of the complex plane. We shall require of the unknown solution  $X = \|x_{ik}\|$  that its characteristic values belong to  $G$  and that their multiplicities be as follows:

$$\begin{aligned} \text{Zeros:} & \quad \lambda_1, \lambda_2, \dots, \\ \text{Multiplicities:} & \quad a_1, a_2, \dots \end{aligned}$$

As in the preceding case, every elementary divisor of  $X$  must have the form

$$(\lambda - \lambda_i)^{p_i} \quad (p_i \leq a_i),$$

and therefore

$$\begin{aligned} X &= T \{ \lambda_{j_1} E^{(p_{j_1})} + H^{(p_{j_1})}, \dots, \lambda_{j_r} E^{(p_{j_r})} + H^{(p_{j_r})} \} T^{-1} \\ (j_1, j_2, \dots, j_r &= 1, 2, \dots; p_{j_1} \leq a_{j_1}, p_{j_2} \leq a_{j_2}, \dots, p_{j_r} \leq a_{j_r}; \\ & p_{j_1} + p_{j_2} + \dots + p_{j_r} = n) \end{aligned} \quad (40)$$

( $T$  is an arbitrary non-singular matrix).

## § 5. Matrix Polynomial Equations

1. Let us consider the equations

$$A_0 X^m + A_1 X^{m-1} + \dots + A_m = O, \quad (41)$$

$$Y^m A_0 + Y^{m-1} A_1 + \dots + A_m = O, \quad (42)$$

where  $A_0, A_1, \dots, A_m$  are given square matrices of order  $n$  and  $X, Y$  are unknown square matrices of the same order. The equation (33) investigated in the preceding section is a very special—one could almost say, trivial—case of (41) and (42) and is obtained by setting  $A_i = a_i E$ , where  $a_i$  is a number and  $i = 1, 2, \dots, m$ .

The following theorem establishes a connection between (41), (42), and (33).

THEOREM 4: Every solution of the matrix equation

$$A_0 X^m + A_1 X^{m-1} + \dots + A_m = O$$

satisfies the scalar equation

$$g(X) = O, \quad (43)$$

where

$$g(\lambda) = |A_0 \lambda^m + A_1 \lambda^{m-1} + \dots + A_m|. \quad (44)$$

The same scalar equation is satisfied by every solution  $Y$  of the matrix equation

$$Y^m A_0 + Y^{m-1} A_1 + \dots + A_m = O.$$

*Proof.* We denote by  $F(\lambda)$  the matrix polynomial

$$F(\lambda) = A_0 \lambda^m + A_1 \lambda^{m-1} + \dots + A_m.$$

Then the equations (41) and (42) can be written as follows (see p. 81) :

$$F(X) = O, \quad \hat{F}(Y) = O.$$

By the generalized Bézout Theorem (Chapter IV, § 3), if  $X$  and  $Y$  are solutions of these equations, the matrix polynomial  $F(\lambda)$  is divisible on the right by  $\lambda E - X$  and on the left by  $\lambda E - Y$  :

$$F(\lambda) = Q(\lambda) (\lambda E - X) = (\lambda E - Y) Q_1(\lambda).$$

Hence

$$g(\lambda) = |F(\lambda)| = |Q(\lambda)| \Delta(\lambda) = |Q_1(\lambda)| \Delta_1(\lambda), \quad (45)$$

where  $\Delta(\lambda) = |\lambda E - X|$  and  $\Delta_1(\lambda) = |\lambda E - Y|$  are the characteristic polynomials of  $X$  and  $Y$ . By the Hamilton-Cayley Theorem (Chapter IV, § 4),

$$\Delta(X) = O, \quad \Delta(Y) = O.$$

Therefore (45) implies that

$$g(X) = g(Y) = O,$$

and the theorem is proved.

Note that the Hamilton-Cayley Theorem is a special case of this theorem. For every square matrix  $A$ , when substituted for  $\lambda$ , satisfies the equation

$$\lambda E - A = O.$$

Therefore, by the theorem just proved,

$$\Delta(A) = O,$$

where  $\Delta(\lambda) = |\lambda E - A|$ .

2. Theorem 4 can be generalized as follows:

THEOREM 5:<sup>5</sup> If  $X_0, X_1, \dots, X_m$  are pairwise permutable square matrices of order  $n$  that satisfy the matrix equation

$$A_0 X_0 + A_1 X_1 + \dots + A_m X_m = O \quad (46)$$

( $A_0, A_1, \dots, A_m$  are given square matrices of order  $n$ ), then the same matrices  $X_0, X_1, \dots, X_m$  satisfy the scalar equation

$$g(X_0, X_1, \dots, X_m) = O, \quad (47)$$

where

$$g(\xi_0, \xi_1, \dots, \xi_m) = |A_0 \xi_0 + A_1 \xi_1 + \dots + A_m \xi_m|. \quad (48)$$

*Proof.* We set<sup>6</sup>

$$F(\xi_0, \xi_1, \dots, \xi_m) = \|f_{ik}(\xi_0, \xi_1, \dots, \xi_m)\|_1^n = A_0 \xi_0 + A_1 \xi_1 + \dots + A_m \xi_m.$$

$\xi_0, \xi_1, \dots, \xi_m$  are scalar variables.

We denote by  $\widehat{F}(\xi_0, \xi_1, \dots, \xi_m) = \|\widehat{f}_{ik}(\xi_0, \xi_1, \dots, \xi_m)\|_1^n$  the adjoint matrix of  $F$  ( $\widehat{f}_{ik}$  is the algebraic complement of  $f_{ki}$  in the determinant  $|F(\xi_0, \xi_1, \dots, \xi_m)| = |f_{ik}|_1^n$  ( $i, k = 1, 2, \dots, n$ )). Then every element  $\widehat{f}_{ik}$  ( $i, k = 1, 2, \dots, n$ ) of  $\widehat{F}$  is a homogeneous polynomial in  $\xi_0, \xi_1, \dots, \xi_m$  of degree  $m-1$ , so that  $\widehat{F}$  can be represented in the form

$$\widehat{F} = \sum_{i_0+i_1+\dots+i_m=n-1} F_{i_0 i_1 \dots i_m} \xi_0^{i_0} \xi_1^{i_1} \dots \xi_m^{i_m},$$

where  $F_{i_0 i_1 \dots i_m}$  are certain constant matrices of order  $n$ .

From the definition of  $\widehat{F}$  there follows the identity

$$\widehat{F}F = g(\xi_0, \xi_1, \dots, \xi_m) E.$$

We write this in the following form:

$$\sum_{i_0+i_1+\dots+i_m=n-1} F_{i_0 i_1 \dots i_m} (A_0 \xi_0 + A_1 \xi_1 + \dots + A_m \xi_m) \xi_0^{i_0} \xi_1^{i_1} \dots \xi_m^{i_m} = g(\xi_0, \xi_1, \dots, \xi_m) E. \quad (49)$$

The transition from the left-hand side of (49) to the right-hand side is accomplished by removing the parentheses and collecting similar terms. In this process we have to permute the variables  $\xi_0, \xi_1, \dots, \xi_m$  among each other, but we do not have to permute the variables  $\xi_0, \xi_1, \dots, \xi_m$  with the matrix coefficients  $A_i$  and  $F_{i_0 i_1 \dots i_m}$ . Therefore the equation (49) is not violated when we substitute for the variables  $\xi_0, \xi_1, \dots, \xi_m$  the pairwise permutable matrices  $X_0, X_1, \dots, X_m$ :

<sup>5</sup> See [318].

<sup>6</sup> The  $f_{ik}(\xi_0, \xi_1, \dots, \xi_m)$  are linear forms in  $\xi_0, \xi_1, \dots, \xi_m$  ( $i, k = 1, 2, \dots, n$ ).



$$\sum_{i_0+i_1+\dots+i_m=n-1} F_{i_0 i_1 \dots i_m} (A_0 X_0 + A_1 X_1 + \dots + A_m X_m) X_0^{i_0} X_1^{i_1} \dots X_m^{i_m} = g(X_0, X_1, \dots, X_m). \quad (50)$$

But, by assumption,

$$A_0 X_0 + A_1 X_1 + \dots + A_m X_m = O.$$

Therefore we find from (50) :

$$g(X_0, X_1, \dots, X_m) = O,$$

and this is what we had to prove.

*Note 1.* Theorem 5 remains valid if (46) is replaced by

$$X_0 A_0 + X_1 A_1 + \dots + X_m A_m = O. \quad (51)$$

For we can apply Theorem 5 to the equation

$$A'_0 X_0 + A'_1 X_1 + \dots + A'_m X_m = O$$

and then go over term by term to the transposed matrices.

*Note 2* Theorem 4 is obtained as a special case of Theorem 5, when we take for  $X_0, X_1, \dots, X_m$

$$X^m, X^{m-1}, \dots, X, E.$$

**3.** We have shown that every solution of (41) satisfies the scalar equation (of degree  $\leq mn$ )

$$g(\lambda) = 0.$$

But the set of matrix solutions of this equation with a given order  $n$  splits into a finite number of classes of similar matrices (see § 4). Therefore all the solutions of (41) have to be looked for among the matrices of the form

$$T_i D_i T_i^{-1} \quad (52)$$

(here  $D_i$  are well-defined matrices; if we wish, we may assume that the  $D_i$  have Jordan normal form.  $T_i$  are arbitrary non-singular matrices of order  $n$ ;  $i = 1, 2, \dots, n$ ). In (41) we substitute for  $X$  the matrix (52) and choose  $T_i$  such that the equation (41) is satisfied. For each  $T_i$  we obtain a linear equation

$$A_0 T_i D_i^m + A_1 T_i D_i^{m-1} + \dots + A_m T_i = O \quad (i = 1, 2, \dots, n). \quad (53)$$

A natural method of finding solutions  $T_i$  of (53) is to replace the matrix equation by a system of linear homogeneous scalar equations in the elements

of the required matrix  $T_i$ . Each non-singular solution  $T_i$  of (53), when substituted in (52), yields a solution of the given equation (41). Similar arguments may be applied to the equation (42).

In the following two sections we shall consider special cases of (41) connected with the extraction of  $m$ -th roots of a matrix.

## § 6. The Extraction of $m$ -th Roots of a Non-Singular Matrix

1. In this section and the following, we deal with the equation

$$X^m = A, \quad (54)$$

where  $A$  is a given matrix and  $X$  an unknown matrix (both of order  $n$ ) and  $m$  is a given positive integer.

In this section we consider the case  $|A| \neq 0$  ( $A$  is non-singular). All the characteristic values of  $A$  are different from zero in this case (since  $|A|$  is the product of these characteristic values).

We denote by

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots, (\lambda - \lambda_u)^{p_u} \quad (55)$$

the elementary divisors of  $A$  and reduce  $A$  to Jordan normal form:<sup>7</sup>

$$A = U \tilde{A} U^{-1} = U \{ \lambda_1 E_1 + H_1, \dots, \lambda_u E_u + H_u \} U^{-1}. \quad (56)$$

Since the characteristic values of the unknown matrix  $X$ , when raised to the  $m$ -th power, give the characteristic values of  $A$ , all the characteristic values of  $X$  are also different from zero. Therefore the derivative of  $f(\lambda) = \lambda^m$  does not vanish on these characteristic values. But then (see Chapter VI, p. 158) the elementary divisors of  $X$  do not 'decompose' when  $X$  is raised to the  $m$ -th power. From what we have said, it follows that the elementary divisors of  $X$  are:

$$(\lambda - \xi_1)^{p_1}, (\lambda - \xi_2)^{p_2}, \dots, (\lambda - \xi_u)^{p_u}. \quad (57)$$

where  $\xi_j^m = \lambda_j$ , i.e.,  $\xi_j$  is one of the  $m$ -th roots of  $\lambda_j$  ( $\xi_j = \sqrt[m]{\lambda_j}$ ;  $j=1, 2, \dots, u$ ).

We now determine  $\sqrt[m]{\lambda_j E_j + H_j}$  in the following way. In the  $\lambda$ -plane we take a circle, with center  $\lambda_j$ , not containing the origin. In this circle we have  $m$  distinct branches of the function  $\sqrt[m]{\lambda}$ . These branches can be distinguished from one another by the value they assume at the center  $\lambda_j$  of the circle. We denote by  $\sqrt[m]{\lambda}$  that branch whose value at  $\lambda_j$  coincides with the characteristic value  $\xi_j$  of the unknown matrix  $X$ , and starting from this branch we define the matrix function  $\sqrt[m]{\lambda_j E_j + H_j}$  by means of the series

<sup>7</sup> Here  $E_j = E(p_j)$  and  $H_j = H(p_j)$  ( $j=1, 2, \dots, u$ ).

$$\sqrt[m]{\lambda_j E_j + H_j} = \lambda_j^{\frac{1}{m}} E_j + \frac{1}{m} \lambda_j^{\frac{1}{m}-1} H_j + \frac{1}{2!} \frac{1}{m} \left( \frac{1}{m} - 1 \right) \lambda_j^{\frac{1}{m}-2} H_j^2 + \dots, \quad (58)$$

which breaks off.

Since the derivative of the function  $\sqrt[m]{\lambda}$  at  $\lambda_j$  is not zero, the matrix (58) has only one elementary divisor  $(\lambda - \xi_j)^{p_j}$ , where  $\xi_j = \sqrt[m]{\lambda}$  (here  $j = 1, 2, 3, \dots, u$ ). Hence it follows that the quasi-diagonal matrix

$$\{\sqrt[m]{\lambda_1 E_1 + H_1}, \sqrt[m]{\lambda_2 E_2 + H_2}, \dots, \sqrt[m]{\lambda_u E_u + H_u}\}$$

has the elementary divisors (57), i.e., the same elementary divisors as the unknown matrix  $X$ . Therefore there exists a non-singular matrix  $T$  ( $|T| \neq 0$ ) such that

$$X = T \{\sqrt[m]{\lambda_1 E_1 + H_1}, \sqrt[m]{\lambda_2 E_2 + H_2}, \dots, \sqrt[m]{\lambda_u E_u + H_u}\} T^{-1}. \quad (59)$$

In order to determine  $T$ , we note that if on both sides of the identity

$$(\sqrt[m]{\lambda})^m = \lambda$$

we substitute the matrix  $\lambda_j E_j + H_j$  ( $j = 1, 2, \dots, u$ ) in place of  $\lambda$ , we obtain:

$$(\sqrt[m]{\lambda_j E_j + H_j})^m = \lambda_j E_j + H_j \quad (j = 1, 2, \dots, u).$$

Now from (54) and (59) it follows that

$$A = T \{\lambda_1 E_1 + H_1, \lambda_2 E_2 + H_2, \dots, \lambda_u E_u + H_u\} T^{-1}. \quad (60)$$

Comparing (56) and (60) we find:

$$T = U X_{\tilde{\lambda}}, \quad (61)$$

where  $X_{\tilde{\lambda}}$  is an arbitrary non-singular matrix permutable with  $\tilde{A}$  (the structure of  $X_{\tilde{\lambda}}$  is described in detail in § 2).

When we substitute in (59) for  $T$  the expression  $U X_{\tilde{\lambda}}$  we obtain a formula that comprises all the solutions of the equation (54):

$$X = U X_{\tilde{\lambda}} \{\sqrt[m]{\lambda_1 E_1 + H_1}, \sqrt[m]{\lambda_2 E_2 + H_2}, \dots, \sqrt[m]{\lambda_u E_u + H_u}\} X_{\tilde{\lambda}}^{-1} U^{-1}. \quad (62)$$

The multivalence of the right-hand side of this formula has a discrete as well as a continuous character: the discrete (in this case finite) character arises from the choice of the distinct branches of the function  $\sqrt[m]{\lambda}$  in the various blocks of the quasi-diagonal matrix (for  $\lambda_i = \lambda_k$  the branches of  $\sqrt[m]{\lambda}$  in the  $j$ -th and  $k$ -th diagonal blocks may even be distinct); the continuous character arises from the arbitrary parameters contained in  $X_{\tilde{\lambda}}$ .

All solutions of (54) will be called  $m$ -th roots of  $A$  and will be denoted by the many-valued symbol  $\sqrt[m]{A}$ . We point out that  $\sqrt[m]{A}$  is, in general, not a function of the matrix  $A$  (i.e., is not representable in the form of a polynomial in  $A$ ).

*Note.* If all the elementary divisors of  $A$  are co-prime in pairs, i.e., if the numbers  $\lambda_1, \lambda_2, \dots, \lambda_u$  are all distinct, then the matrix  $X_{\tilde{A}}$  has quasi-diagonal form

$$X_{\tilde{A}} = \{X_1, X_2, \dots, X_u\},$$

where  $X_j$  is permutable with  $\lambda_j E_j + H_j$  and therefore permutable with every function of  $\lambda_j E_j + H_j$  and, in particular, with  $\sqrt[m]{\lambda_j E_j + H_j}$  ( $j = 1, 2, \dots, u$ ). Therefore in this case (62) assumes the form

$$X = U \{ \sqrt[m]{\lambda_1 E_1 + H_1}, \sqrt[m]{\lambda_2 E_2 + H_2}, \dots, \sqrt[m]{\lambda_u E_u + H_u} \} U^{-1}.$$

Thus, if the elementary divisors of  $A$  are co-prime in pairs, then in the formula for  $X = \sqrt[m]{A}$  only a discrete multivalence occurs. In this case every value of  $\sqrt[m]{A}$  can be represented as a polynomial in  $A$ .

2. *Example.* Suppose it is required to find all square roots of

$$A = \begin{vmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix},$$

i.e., all solutions of the equation

$$X^2 = A.$$

In this case  $A$  has already the Jordan normal form. Therefore in (62) we can set  $A = \tilde{A}$ ,  $U = E$ . The matrix  $X_{\tilde{A}}$  in this case looks as follows:

$$X_{\tilde{A}} = \begin{vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{vmatrix},$$

where  $a, b, c, d$ , and  $e$  are arbitrary parameters.

The formula (62), which gives all the required solutions  $X$ , now assumes the following form:

$$X = \begin{vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{vmatrix} \begin{vmatrix} \varepsilon & \frac{\varepsilon}{2} & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & \eta \end{vmatrix} \begin{vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & e \end{vmatrix}^{-1} \quad (\varepsilon^2 = \eta^2 = 1). \quad (63)$$

Without changing  $X$  we may multiply  $X_{\tilde{A}}$  in (62) by a scalar so that  $|X_{\tilde{A}}| = 1$ . Then this leads to the equation  $a^2 e = 1$ ; and hence  $e = a^{-2}$ .

Let us compute the elements of  $X_{\tilde{A}}^{-1}$ . For this purpose we write down the linear transformation with the matrix coefficients of  $X_{\tilde{A}}$ :

$$\begin{aligned}y_1 &= ax_1 + bx_2 + cx_3, \\y_2 &= ax_2, \\y_3 &= dx_3 + a^{-2}x_2.\end{aligned}$$

We solve this system of equations with respect to  $x_1, x_2, x_3$ . Then we obtain the transformation with the inverse matrix  $X_{\tilde{A}}^{-1}$ :

$$\begin{aligned}x_1 &= a^{-1}y_1 - (a^{-2}b - cd)y_2 - acy_3, \\x_2 &= a^{-1}y_2, \\x_3 &= -a dy_2 + a^2 y_3.\end{aligned}$$

Hence we find:

$$X_{\tilde{A}}^{-1} = \begin{vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & a^{-2} \end{vmatrix}^{-1} = \begin{vmatrix} a^{-1} & cd - a^{-2}b & -ac \\ 0 & a^{-1} & 0 \\ 0 & -ad & a^2 \end{vmatrix}.$$

The formula (63) yields:

$$\begin{aligned}X &= \begin{vmatrix} \varepsilon & (\varepsilon - \eta) a cd + \frac{\varepsilon}{2} a^2 c (\eta - \varepsilon) \\ 0 & \varepsilon & 0 \\ 0 & (\varepsilon - \eta) da^{-1} & \eta \end{vmatrix} \\ &= \begin{vmatrix} \varepsilon & (\varepsilon - \eta) vw + \frac{\varepsilon}{2} (\eta - \varepsilon) v \\ 0 & \varepsilon & 0 \\ 0 & (\varepsilon - \eta) w & \eta \end{vmatrix} \quad (v = a^2 c, w = a^{-1} d). \quad (64)\end{aligned}$$

The solution  $X$  depends on two arbitrary parameters  $u$  and  $w$  and two arbitrary signs  $\varepsilon$  and  $\eta$ .

## § 7. The Extraction of $m$ -th Roots of a Singular Matrix

1. We pass on to the discussion of the case where  $|A| = 0$  ( $A$  is a singular matrix).

As in the first case, we reduce  $A$  to the Jordan normal form:

$$A = U \{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)}; H^{(q_1)}, H^{(q_1)}, \dots, H^{(q_r)} \} U^{-1}; \quad (65)$$

here we have denoted by  $(\lambda - \lambda_1)^{p_1}, \dots, (\lambda - \lambda_u)^{p_u}$  the elementary divisors of  $A$  that correspond to non-zero characteristic values, and by  $\lambda^{q_1}, \lambda^{q_2}, \dots, \lambda^{q_r}$  the elementary divisors with characteristic value zero.

Then

$$A = U \{A_1, A_2\} U^{-1}, \quad (66)$$

where

$$A_1 = \{\lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)}\}, A_2 = \{H^{(q_1)}, H^{(q_2)}, \dots, H^{(q_r)}\}. \quad (67)$$

Note that  $A_1$  is a non-singular matrix ( $|A_1| \neq 0$ ) and  $A_2$  a nilpotent matrix with index of nilpotency  $\mu = \max(q_1, q_2, \dots, q_r)$  ( $A_2^\mu = O$ ).

The original equation (54) implies that  $A$  commutes with the unknown matrix  $X$  and therefore the similar matrices

$$U^{-1}AU = \{A_1, A_2\} \text{ and } U^{-1}XU \quad (68)$$

also commute.

As we have shown in § 2 (Theorem 3), from the permutability of the matrices (68) and the fact that  $A_1$  and  $A_2$  do not have characteristic values in common, it follows that the second matrix in (68) has a corresponding quasi-diagonal form

$$U^{-1}XU = \{X_1, X_2\}. \quad (69)$$

When we replace the matrices  $A$  and  $X$  in (54) by the similar matrices

$$\{A_1, A_2\} \text{ and } \{X_1, X_2\},$$

we replace (54) by two equations:

$$X_1^m = A_1, \quad (70)$$

$$X_2^m = A_2. \quad (71)$$

Since  $|A_1| \neq 0$ , the results of the preceding section are applicable to (70). Therefore we find  $X_1$  by the formula (62):

$$X_1 = X_{A_1} \{ \sqrt[m]{\lambda_1 E^{(p_1)} + H^{(p_1)}}, \dots, \sqrt[m]{\lambda_u E^{(p_u)} + H^{(p_u)}} \} X_{A_1}^{-1}. \quad (72)$$

Thus it remains to consider the equation (71), i.e., to find all  $m$ -th roots of the nilpotent matrix  $A_2$ , which already has the Jordan normal form

$$A_2 = \{H^{(q_1)}, H^{(q_2)}, \dots, H^{(q_r)}\}; \quad (73)$$

$\mu = \max(q_1, q_2, \dots, q_r)$  is the index of nilpotency of  $A_2$ .

From  $A_2^\mu = O$  and (71) we find

$$X_2^{m\mu} = O.$$

The last equation shows that the required matrix  $X_2$  is also nilpotent with an index of nilpotency  $\nu$ , where  $m(\mu - 1) < \nu \leq m\mu$ . We reduce  $X_2$  to the Jordan form:

$$X_2 = T \{H^{(v_1)}, H^{(v_2)}, \dots, H^{(v_r)}\} T^{-1} \quad (74)$$

$$(v_1, v_2, \dots, v_r \leq r).$$

Now we raise both sides of (74) to the  $m$ -th power. We obtain:

$$A_2 = X_2^m = T \{[H^{(v_1)}]^m, [H^{(v_2)}]^m, \dots, [H^{(v_r)}]^m\} T^{-1}. \quad (75)$$

2. Let us now clarify the question of what elementary divisors the matrix  $[H^{(v)}]^m$  has.<sup>5</sup> We denote by  $H$  the linear operator given by  $H^{(v)}$  in a  $r$ -dimensional vector space with the basis  $e_1, e_2, \dots, e_r$ . Then from the form of the matrix  $H^{(v)}$  (in  $H^{(v)}$  all the elements of the first superdiagonal are equal to 1 and all the remaining elements are 0) it follows that

$$He_1 = o, He_2 = e_1, \dots, He_r = e_{r-1}. \quad (76)$$

These equations show that the vectors  $e_1, e_2, \dots, e_r$  form a Jordan chain for  $H$ , corresponding to the elementary divisor  $\lambda^r$ .

We write (76) as follows:

$$He_j = e_{j-1} \quad (j=1, 2, \dots, r; e_0 = o).$$

Obviously,

$$H^m e_j = e_{j-m} \quad (j=1, 2, \dots, r; e_0 = e_{-1} = \dots = e_{-m+1} = o). \quad (77)$$

We express  $v$  in the form

$$v = km + r \quad (r < m),$$

where  $k$  and  $r$  are non-negative integers. We arrange the basis vectors  $e_1, e_2, \dots, e_r$  in the following way:

$$\begin{array}{ccccccc} e_1, & e_2, & \dots, & e_m, & & & \\ e_{m+1}, & e_{m+2}, & \dots, & e_{2m}, & & & \\ \dots & \dots & \dots & \dots & \dots & \dots & \\ e_{(k-1)m+1}, & e_{(k-1)m+2}, & \dots, & e_{km}, & & & \\ e_{km+1}, & \dots, & e_{km+r}. & & & & \end{array} \quad (78)$$

This table has  $m$  columns: the first  $r$  columns contain  $k+1$  vectors each, the remaining ones  $k$  vectors. The equation (77) shows that the vectors of each column form a Jordan chain with respect to the operator  $H^m$ . If instead

<sup>5</sup> This question is answered by Theorem 9 of Chapter VI (p. 158). Here we are compelled to use another method of investigating the problem, because we have to find not only the elementary divisor of the matrix  $[H^{(v)}]^m$ , but also a matrix  $p_{v,m}$  transforming  $[H^{(v)}]^m$  into Jordan form.

of numbering the vectors (78) by rows we number them by columns, we obtain a new basis in which the matrix of the operator  $H^m$  has the following Jordan normal form:<sup>9</sup>

$$\underbrace{\{H^{(k+1)}, \dots, H^{(k+1)}\}}_r, \underbrace{\{H^{(k)}, \dots, H^{(k)}\}}_{m-r};$$

and therefore

$$[H^{(v)}]^m = P_{v,m} \underbrace{\{H^{(k+1)}, \dots, H^{(k+1)}\}}_r \underbrace{\{H^{(k)}, \dots, H^{(k)}\}}_{m-r} P_{v,m}^{-1}, \quad (79)$$

where the matrix  $P_{v,m}$  (describing the transition from the one basis to the other) has the following form (see Chapter III, § 4):

$$P_{v,m} = \left\| \begin{array}{cccccc} & \overbrace{\phantom{0 \dots 0}}^m & & & & \\ 1 & 0 & \dots & 0 & 0 & \dots \\ 0 & 0 & \dots & 0 & 1 & \dots \\ \vdots & \vdots & & \vdots & \vdots & \\ 0 & 0 & \dots & 0 & & \\ 0 & 1 & \dots & 0 & & \\ \vdots & \vdots & & \vdots & \vdots & \end{array} \right\| m. \quad (80)$$

The matrix  $H^{(v)}$  has the single elementary divisor  $\lambda^v$ . When  $H^{(v)}$  is raised to the  $m$ -th power, this elementary divisor 'falls apart.' As (79) shows,  $[H^{(v)}]^m$  has the elementary divisors:

$$\underbrace{\lambda^{k+1}, \dots, \lambda^{k+1}}_r, \underbrace{\lambda^k, \dots, \lambda^k}_{m-r}$$

Turning now to (75), we set:

$$v_i = k_i m + r_i \quad (0 \leq r_i < m, k_i \geq 0; \quad i = 1, 2, \dots, s). \quad (81)$$

Then, by (79), equation (75) can be written as follows:

$$A_2 = X_2^m = T P \underbrace{\{H^{(k_1+1)}, \dots, H^{(k_1+1)}\}}_{r_1} \underbrace{\{H^{(k_1)}, \dots, H^{(k_1)}\}}_{m-r_1} \underbrace{\{H^{(k_2+1)}, \dots, H^{(k_2+1)}\}}_{r_2} \underbrace{\{H^{(k_2)}, \dots, H^{(k_2)}\}}_{m-r_2} P^{-1} T^{-1}. \quad (82)$$

where

$$P = \{P_{v_1, m}, P_{v_2, m}, \dots, P_{v_s, m}\}$$

<sup>9</sup> In the case  $k = 0$ , the blocks  $\underbrace{H^{(k)}, \dots, H^{(k)}}_{m-r}$  are absent, and the matrix has the form  $\underbrace{\{H^{(1)}, \dots, H^{(1)}\}}_r$ .



(Comparing (82) with (73), we see that the blocks

$$H^{(k_1+1)}, \dots, H^{(k_1+1)}, H^{(k_1)}, \dots, H^{(k_1)}, H^{(k_1+1)}, \dots, H^{(k_1+1)}, \dots \quad (83)$$

must coincide, apart from the order, with the blocks

$$H^{(q_1)}, H^{(q_2)}, \dots, H^{(q_s)} \quad (84)$$

3. Let us call a system of elementary divisors  $\lambda^{v_1}, \lambda^{v_2}, \dots, \lambda^{v_s}$  *admissible* for  $X_2$  if after raising of the matrix to the  $m$ -th power these elementary divisors split and generate the given system of elementary divisors of  $A_2$ :  $\lambda^{q_1}, \lambda^{q_2}, \lambda^{q_3}, \dots, \lambda^{q_t}$ . The number of admissible systems of elementary divisors is always finite, because

$$\max(v_1, v_2, \dots, v_s) \leq m\mu, \quad v_1 + v_2 + \dots + v_s = n_2 \quad (85)$$

( $n_2$  is the order of  $A_2$ ).

In every concrete case the admissible systems of elementary divisors for  $X_2$  can easily be determined by a finite number of trials.

Let us show that for each admissible system of elementary divisors  $\lambda^{v_1}, \lambda^{v_2}, \dots, \lambda^{v_s}$  form a corresponding solution of (71) and let us determine all these solutions. In this case there exists a transforming matrix  $Q$  such that

$$\{H^{(k_1+1)}, \dots, H^{(k_1+1)}, H^{(k_1)}, \dots, H^{(k_1)}, H^{(k_1+1)}, \dots\} = Q^{-1} A_2 Q. \quad (86)$$

The matrix  $Q$  describes the permutation of the blocks in the quasi-diagonal matrix that brings about the proper renumbering of the basis vectors. Therefore  $Q$  can be regarded as known. Using (86), we obtain from (82):

$$A_2 = TPQ^{-1}A_2QP^{-1}T^{-1}.$$

Hence

$$TPQ^{-1} = X_{A_2},$$

or

$$T = X_{A_2} Q P^{-1}, \quad (87)$$

where  $X_{A_2}$  is an arbitrary matrix that commutes with  $A_2$ .

Substituting (87) for  $T$  in (74), we have

$$X_2 = X_{A_2} Q P^{-1} \{H^{(v_1)}, H^{(v_2)}, \dots, H^{(v_s)}\} P Q^{-1} X_{A_2}^{-1}. \quad (88)$$

From (69), (72), and (88) we obtain a general formula which comprises all the solutions:

$$X = U \{X_{A_2}, X_{A_2} Q P^{-1}\} \left\{ \sqrt[m]{\lambda_1 E^{(p_1)} + H^{(p_1)}}, \dots, \sqrt[m]{\lambda_r E^{(p_r)} + H^{(p_r)}}, \right. \\ \left. H^{(v_1)}, \dots, H^{(v_s)} \right\} \cdot \{X_{A_2}^{-1}, P Q^{-1} X_{A_2}^{-1}\} U^{-1}. \quad (89)$$

We draw the reader's attention to the fact that the  $m$ -th root of a singular matrix does not always exist. Its existence is bound up with the existence of a system of admissible elementary divisors for  $X_2$ .

It is easy to see, for example, that the equation

$$X^m = H^{(p)}$$

has no solution for  $m > 1$ ,  $p > 1$ .

*Example.* Suppose it is required to extract the square root of

$$A = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix},$$

i.e., to find all the solutions of the equation

$$X^2 = A.$$

In this case,  $A = A_2$ ,  $X = X_2$ ,  $m = 2$ ,  $t = 2$ ,  $q_1 = 2$ , and  $q_2 = 1$ . The matrix  $X$  can only have the one elementary divisor  $\lambda^3$ . Therefore  $s = 1$ ,  $v_1 = 3$ ,  $k_1 = 1$ ,  $r_1 = 1$  and (see (80))

$$P = P_{2,1} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{vmatrix} = P^{-1}, Q = E.$$

Moreover, as in the example on page 233, in (88) we may set:

$$X_{A_1} = \begin{vmatrix} a & b & c \\ 0 & a & 0 \\ 0 & d & a^{-2} \end{vmatrix}, \quad X_{A_1}^{-1} = \begin{vmatrix} a^{-1} & cd - a^{-2}b & -ac \\ 0 & a^{-1} & 0 \\ 0 & -ad & a^3 \end{vmatrix}.$$

From this formula we obtain

$$X = X_1 = X_{A_1} P^{-1} H^{(3)} P X_{A_1}^{-1} = \begin{vmatrix} 0 & \alpha & \beta \\ 0 & 0 & 0 \\ 0 & \beta^{-1} & 0 \end{vmatrix},$$

where  $\alpha = ca^{-1} - a^2d$  and  $\beta = a^3$  are arbitrary parameters.

## § 8. The Logarithm of a Matrix

1. We consider the matrix equation

$$e^X = A. \tag{90}$$

All the solutions of this equation are called (natural) *logarithms* of  $A$  and are denoted by  $\ln A$ .

The characteristic values  $\lambda_j$  of  $A$  are connected with the characteristic values  $\xi_j$  of  $X$  by the formula  $\lambda_j = e^{\xi_j}$ ; therefore, if the equation (90) has a solution, then all the characteristic values of  $A$  are different from zero, and  $A$  is non-singular ( $|A| \neq 0$ ). Thus, the condition  $|A| \neq 0$  is necessary for the existence of solutions of the equation (90). Below, we shall see that this condition is also sufficient.

Suppose, then, that  $|A| \neq 0$ . We write down the elementary divisors of  $A$ :

$$(\lambda - \lambda_1)^{p_1}, (\lambda - \lambda_2)^{p_2}, \dots, (\lambda - \lambda_u)^{p_u} \\ (\lambda_1 \lambda_2 \dots \lambda_u \neq 0, \quad p_1 + p_2 + \dots + p_u = n). \quad (91)$$

Corresponding to these elementary divisors we reduce  $A$  to the Jordan normal form:

$$A = U \tilde{A} U^{-1} \\ = U \{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \lambda_2 E^{(p_2)} + H^{(p_2)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)} \} U^{-1}. \quad (92)$$

Since the derivative of the function  $e^\xi$  is different from zero for all values of  $\xi$ , we know (see Chapter VI, p. 158) that in the transition from  $X$  to  $A = e^X$  the elementary divisors do not split, so that  $X$  has the elementary divisors

$$(\lambda - \xi_1)^{p_1}, (\lambda - \xi_2)^{p_2}, \dots, (\lambda - \xi_u)^{p_u}, \quad (93)$$

where  $e^{\xi_j} = \lambda_j$  ( $j = 1, 2, \dots, u$ ), i.e.,  $\xi_j$  is one of the values of  $\ln \lambda_j$  ( $j = 1, 2, 3, \dots, u$ ).

In the plane of the complex variable  $\lambda$  we draw a circle with center at  $\lambda_j$  and with radius less than  $|\lambda_j|$  and we denote by  $f_j(\lambda) = \ln \lambda$  that branch of the function  $\ln \lambda$  in this circle which at  $\lambda_j$  assumes the value equal to the characteristic value  $\xi_j$  of  $X$  ( $j = 1, 2, \dots, u$ ). After this, we set:

$$\ln(\lambda_j E^{(p_j)} + H^{(p_j)}) = f_j(\lambda_j E^{(p_j)} + H^{(p_j)}) = \ln \lambda_j E^{(p_j)} + \lambda_j^{-1} H^{(p_j)} + \dots \quad (94)$$

Since the derivative of  $\ln \lambda$  vanishes nowhere (in the finite part of the  $\lambda$ -plane), the matrix (94) has only the one elementary divisor  $(\lambda - \xi_j)^{p_j}$ . Therefore the quasi-diagonal matrix

$$\{ \ln(\lambda_1 E^{(p_1)} + H^{(p_1)}), \ln(\lambda_2 E^{(p_2)} + H^{(p_2)}), \dots, \ln(\lambda_u E^{(p_u)} + H^{(p_u)}) \} \quad (95)$$

has the same elementary divisors as the unknown matrix  $X$ . Therefore there exists a matrix  $T$  ( $|T| \neq 0$ ) such that

$$\mathbf{X} = T \{ \ln(\lambda_1 E^{(p_1)} + H^{(p_1)}), \dots, \ln(\lambda_u E^{(p_u)} + H^{(p_u)}) \} T^{-1}. \quad (96)$$

In order to determine  $T$ , we note that

$$A = e^{\mathbf{X}} = T \{ \lambda_1 E^{(p_1)} + H^{(p_1)}, \dots, \lambda_u E^{(p_u)} + H^{(p_u)} \} T^{-1}. \quad (97)$$

Comparing (97) and (92), we find:

$$T = UX_{\tilde{A}}, \quad (98)$$

where  $X_{\tilde{A}}$  is an arbitrary matrix that commutes with  $\tilde{A}$ . Substituting the expression for  $T$  from (98) into (96), we obtain a general formula that comprises all the logarithms of the matrix:

$$\mathbf{X} = UX_{\tilde{A}} \{ \ln(\lambda_1 E^{(p_1)} + H^{(p_1)}), \dots, \ln(\lambda_u E^{(p_u)} + H^{(p_u)}) \} X_{\tilde{A}}^{-1} U^{-1}. \quad (99)$$

*Note.* If all the elementary divisors of  $A$  are co-prime, then on the right-hand side of (99) the factors  $X_{\tilde{A}}$  and  $X_{\tilde{A}}^{-1}$  can be omitted (see a similar remark on p. 233).

## CHAPTER IX

### LINEAR OPERATORS IN A UNITARY SPACE

#### § 1. General Considerations

In Chapters III and VII we studied linear operators in an arbitrary  $n$ -dimensional vector space. All the bases of such a space are of equal standing. To a given linear operator there corresponds in each basis a certain matrix. The matrices corresponding to one and the same operator in the various bases are similar. Thus, the study of linear operators in an  $n$ -dimensional vector space enables us to bring out those properties of matrices that are inherent in an entire class of similar matrices.

At the beginning of this chapter we shall introduce a metric into an  $n$ -dimensional space by assigning in a special way to each pair of vectors a certain number, the 'scalar product' of the two vectors. By means of the scalar product we shall define the 'length' of a vector and the cosine of the 'angle' between two vectors. This metrization leads to a unitary space if the ground field  $\mathbb{F}$  is the field of all complex numbers and to a euclidean space if  $\mathbb{F}$  is the field of all real numbers.

In the present chapter we shall study the properties of linear operators that are connected with the metric of the space. All the bases of the space are by no means of equal standing with respect to the metric. However, this does hold true of all *orthonormal* bases. The transition from one orthonormal basis to another in a unitary space is brought about by means of a special—namely, unitary—transformation (in a euclidean space, an orthogonal transformation). Therefore all the matrices that correspond to one and the same linear operator in two distinct bases of a unitary (euclidean) space are unitarily (orthogonally) similar. Thus, by studying linear operators in an  $n$ -dimensional metrized space we study the properties of matrices that remain invariant under transition from a given matrix to a unitarily—or orthogonally—similar one. This will lead in a natural way to the investigation of properties of special classes of matrices (normal, hermitian, unitary, symmetric, skew-symmetric, orthogonal matrices).

## § 2. Metrization of a Space

1. We consider a vector space  $\mathbf{R}$  over the field of complex numbers. To every pair of vectors  $\mathbf{x}$  and  $\mathbf{y}$  of  $\mathbf{R}$  given in a definite order let a certain complex number be assigned, the so-called *scalar product*, or *inner product*, of the vectors, denoted by  $(\mathbf{x}\mathbf{y})$  or  $(\mathbf{x}, \mathbf{y})$ . Suppose further that the 'scalar multiplication' has the following properties:

For arbitrary vectors  $\mathbf{x}, \mathbf{y}, \mathbf{z}$  of  $\mathbf{R}$  and an arbitrary complex number  $\alpha$ , let<sup>1</sup>

$$\left. \begin{aligned} 1. \quad & (\mathbf{x}\mathbf{y}) = \overline{(\mathbf{y}\mathbf{x})}, \\ 2. \quad & (\alpha\mathbf{x}, \mathbf{y}) = \alpha(\mathbf{x}\mathbf{y}), \\ 3. \quad & (\mathbf{x} + \mathbf{y}, \mathbf{z}) = (\mathbf{x}\mathbf{z}) + (\mathbf{y}\mathbf{z}). \end{aligned} \right\} \quad (1)$$

Then we shall say that a *hermitian metric* is introduced in  $\mathbf{R}$ .

Note that 1., 2., and 3. have the following consequences for arbitrary  $\mathbf{x}, \mathbf{y}, \mathbf{z}$  in  $\mathbf{R}$ :

$$\begin{aligned} 2'. \quad & (\mathbf{x}, \alpha\mathbf{y}) = \bar{\alpha}(\mathbf{x}\mathbf{y}), \\ 3'. \quad & (\mathbf{x}, \mathbf{y} + \mathbf{z}) = (\mathbf{x}\mathbf{y}) + (\mathbf{x}\mathbf{z}). \end{aligned}$$

From 1. we deduce that for every vector  $\mathbf{x}$  the scalar product  $(\mathbf{x}\mathbf{x})$  is a real number. This number is called the *norm* of  $\mathbf{x}$  and is denoted by  $N\mathbf{x}$ :  $N\mathbf{x} = (\mathbf{x}, \mathbf{x})$ .

If for every vector  $\mathbf{x}$  of  $\mathbf{R}$

$$4. \quad N\mathbf{x} = (\mathbf{x}\mathbf{x}) \geq 0, \quad (2)$$

then the hermitian metric is called *positive semi-definite*. And if, moreover,

$$5. \quad N\mathbf{x} = (\mathbf{x}\mathbf{x}) > 0 \text{ for } \mathbf{x} \neq \mathbf{o}, \quad (3)$$

then the hermitian metric is called *positive definite*.

**DEFINITION 1:** A vector space  $\mathbf{R}$  with a positive-definite hermitian metric will be called a *unitary space*.<sup>2</sup>

In this chapter we shall consider finite-dimensional unitary spaces.<sup>3</sup>

By the *length* of the vector  $\mathbf{x}$  we mean<sup>4</sup>  $+\sqrt{N\mathbf{x}} = +\sqrt{(\mathbf{x}, \mathbf{x})} = |\mathbf{x}|$ . From 2. and 5. it follows that every vector other than the null vector has a positive

<sup>1</sup> A number with a bar over it denotes the complex conjugate of the number.

<sup>2</sup> The study of  $n$ -dimensional vector spaces with an arbitrary (not positive-definite) metric is taken up in the paper [319].

<sup>3</sup> In §§ 2-7 of this chapter, wherever it is not expressly stated that the space is finite-dimensional, all the arguments remain valid for infinite-dimensional spaces.

<sup>4</sup> The symbol  $+\sqrt{\phantom{x}}$  denotes the non-negative (arithmetical) value of the root.

length and that the null vector has length 0. A vector  $\mathbf{x}$  is called *normalized* (or is said to be a *unit vector*) if  $|\mathbf{x}| = 1$ . To normalize an arbitrary vector  $\mathbf{x} \neq \mathbf{o}$  it is sufficient to multiply it by any complex number  $\lambda$  for which  $|\lambda| = \frac{1}{|\mathbf{x}|}$ .

By analogy with the ordinary three-dimensional vector spaces, two vectors  $\mathbf{x}$  and  $\mathbf{y}$  are called *orthogonal* (in symbols:  $\mathbf{x} \perp \mathbf{y}$ ) if  $(\mathbf{x}\mathbf{y}) = 0$ . In this case it follows from 1., 3., and 3'. that

$$N(\mathbf{x} + \mathbf{y}) = (\mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y}) = (\mathbf{x}\mathbf{x}) + (\mathbf{y}\mathbf{y}) = N\mathbf{x} + N\mathbf{y},$$

i.e. (the theorem of Pythagoras!),

$$|\mathbf{x} + \mathbf{y}|^2 = |\mathbf{x}|^2 + |\mathbf{y}|^2 \quad (\mathbf{x} \perp \mathbf{y}).$$

Let  $\mathbf{R}$  be a unitary space of finite dimension  $n$ . We consider an arbitrary basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  of  $\mathbf{R}$ . Let us denote by  $x_i$  and  $y_i$  ( $i = 1, 2, \dots, n$ ) the coordinates of the vectors  $\mathbf{x}$  and  $\mathbf{y}$  in this basis:

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i, \quad \mathbf{y} = \sum_{i=1}^n y_i \mathbf{e}_i.$$

Then by 2., 3., 2', and 3',

$$(\mathbf{x}\mathbf{y}) = \sum_{i,k=1}^n h_{ik} x_i \bar{y}_k, \quad (4)$$

where

$$h_{ik} = (\mathbf{e}_i \mathbf{e}_k) \quad (i, k = 1, 2, \dots, n). \quad (5)$$

In particular,

$$N\mathbf{x} = (\mathbf{x}\mathbf{x}) = \sum_{i,k=1}^n h_{ik} x_i \bar{x}_k. \quad (6)$$

From 1. and (5) we deduce

$$h_{ki} = \bar{h}_{ik} \quad (i, k = 1, 2, \dots, n). \quad (7)$$

2. A form  $\sum_{i,k=1}^n \bar{h}_{ik} x_i \bar{x}_k$ , where  $h_{ki} = h_{ik}$  ( $i, k = 1, 2, \dots, n$ ) is called *hermitian*.<sup>5</sup> Thus, the norm of a vector, i.e., the square of its length, is a hermitian form in its coordinates. Hence the name 'hermitian metric.' The form on the right-hand side of (6) is, by 4., *non-negative*:

$$\sum_{i,k=1}^n h_{ik} x_i \bar{x}_k \geq 0 \quad (8)$$

for all values of the variables  $x_1, x_2, \dots, x_n$ . By the additional condition 5., the form is in fact *positive definite*, i.e., the equality sign in (8) only holds when all the  $x_i$  are zero ( $i = 1, 2, \dots, n$ ).

<sup>5</sup> In accordance with this, the expression on the right-hand side of (4) is called a *hermitian bilinear form* (in  $x_1, x_2, \dots, x_n$  and  $y_1, y_2, \dots, y_n$ ).

DEFINITION 2: A system of vectors  $e_1, e_2, \dots, e_m$  is called *orthonormal* if

$$(e_i, e_k) = \delta_{ik} = \begin{cases} 0, & \text{for } i \neq k, \\ 1, & \text{for } i = k \end{cases} \quad (i, k = 1, 2, \dots, m). \quad (9)$$

When  $m = n$ , where  $n$  is the dimension of the space, we obtain an *orthonormal basis* of the space.

In § 7 we shall prove that every  $n$ -dimensional space has an *orthonormal basis*.

Let  $x_i$  and  $y_i$  ( $i = 1, 2, \dots, n$ ) be the coordinates of  $\mathbf{x}$  and  $\mathbf{y}$  in an orthonormal basis. Then by (4), (5), and (9)

$$\left. \begin{aligned} (\mathbf{x}\mathbf{y}) &= \sum_{i=1}^n x_i \bar{y}_i, \\ \mathbf{N}\mathbf{x} = (\mathbf{x}\mathbf{x}) &= \sum_{i=1}^n |x_i|^2. \end{aligned} \right\} \quad (10)$$

Let us take an arbitrary fixed basis in an  $n$ -dimensional space  $\mathbf{R}$ . In this basis every metrization of the space is connected with a certain positive-definite hermitian form  $\sum_{i,k=1}^n h_{ik} x_i \bar{x}_k$ ; and conversely, by (4), every such form determines a certain positive-definite hermitian metric in  $\mathbf{R}$ . However, these metrics do not all give essentially different unitary  $n$ -dimensional spaces. For let us take two such metrics with the respective scalar products  $(\mathbf{x}\mathbf{y})$  and  $(\mathbf{x}\mathbf{y})'$ . We determine orthonormal bases in  $\mathbf{R}$  with respect to these metrics:  $e_i$  and  $e'_i$  ( $i = 1, 2, \dots, n$ ). Let the vector  $\mathbf{x}$  in  $\mathbf{R}$  be mapped onto the vector  $\mathbf{x}'$  in  $\mathbf{R}$ , where  $\mathbf{x}'$  is the vector whose coordinates in the basis  $e'_i$  are the same as the coordinates of  $\mathbf{x}$  in the basis  $e_i$  ( $i = 1, 2, \dots, n$ ). ( $\mathbf{x} \rightarrow \mathbf{x}'$ .) This mapping is *affine*.<sup>6</sup> Moreover, by (10),

$$(\mathbf{x}\mathbf{y}) = (\mathbf{x}'\mathbf{y})'.$$

Therefore: To within an affine transformation of the space all positive definite hermitian metrizations of an  $n$ -dimensional vector space coincide.

If the field  $\mathbb{F}$  is the field of real numbers, then a metric satisfying the postulates 1., 2., 3., 4., and 5. is called *euclidean*.

DEFINITION 3: A vector space  $\mathbf{R}$  over the field of real numbers with a positive euclidean metric is called a *euclidean space*.

If  $x_i$  and  $y_i$  ( $i = 1, 2, \dots, n$ ) are the coordinates of the vectors  $\mathbf{x}$  and  $\mathbf{y}$  in some basis  $e_1, e_2, \dots, e_n$  of an  $n$ -dimensional euclidean space, then

<sup>6</sup> I.e., the operator  $A$  that maps the vector  $\mathbf{x}$  of  $\mathbf{R}$  onto the vector  $\mathbf{x}'$  of  $\mathbf{R}'$  is linear and non-singular.



<sup>8</sup> In the case of a euclidean space,  $c_1, c_2, \dots, c_m$  are real numbers.



$$(f, g) = \int_{\alpha}^{\beta} f(t) \overline{g(t)} dt.$$

Then Gram's criterion (Theorem 1) applied to the given function yields the required condition:

$$\begin{vmatrix} \int_{\alpha}^{\beta} f_1(t) \overline{f_1(t)} dt & \dots & \int_{\alpha}^{\beta} f_1(t) \overline{f_n(t)} dt \\ \dots & \dots & \dots \\ \int_{\alpha}^{\beta} f_n(t) \overline{f_1(t)} dt & \dots & \int_{\alpha}^{\beta} f_n(t) \overline{f_n(t)} dt \end{vmatrix} = 0.$$

#### § 4. Orthogonal Projection

1. Let  $\mathbf{x}$  be an arbitrary vector in a unitary or euclidean space  $\mathbf{R}$  and  $\mathbf{S}$  an  $m$ -dimensional subspace with a basis  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ . We shall show that  $\mathbf{x}$  can be represented (and moreover, represented uniquely) in the form

$$\mathbf{x} = \mathbf{x}_S + \mathbf{x}_N, \quad (15)$$

where

$$\mathbf{x}_S \in \mathbf{S} \text{ and } \mathbf{x}_N \perp \mathbf{S}$$

(the symbol  $\perp$  denotes orthogonality of vectors; orthogonality to a subspace means orthogonality to every vector of the subspace);  $\mathbf{x}_S$  is the orthogonal projection of  $\mathbf{x}$  onto  $\mathbf{S}$ ,  $\mathbf{x}_N$  the projecting vector.

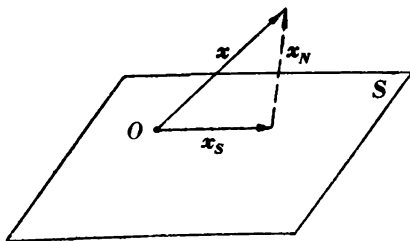


Fig. 5

*Example.* Let  $\mathbf{R}$  be a three-dimensional euclidean vector space and  $m = 2$ . Let all vectors originate at a fixed point  $O$ . Then  $\mathbf{S}$  is a plane passing through  $O$ ;  $\mathbf{x}_S$  is the orthogonal projection of  $\mathbf{x}$  onto the plane  $\mathbf{S}$ ;  $\mathbf{x}_N$  is the perpendicular dropped from the endpoint of  $\mathbf{x}$  onto the plane  $\mathbf{S}$  (Fig. 5); and  $h = |\mathbf{x}_N|$  is the distance of the endpoint of  $\mathbf{x}$  from  $\mathbf{S}$ .

To establish the decomposition (15), we represent the required  $\mathbf{x}_S$  in the form

$$\mathbf{x}_S = c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + \dots + c_m \mathbf{x}_m, \quad (16)$$

where  $c_1, c_2, \dots, c_m$  are complex numbers.\*

\* In the case of a euclidean space,  $c_1, c_2, \dots, c_m$  are real numbers.

To determine these numbers we shall start from the relations

$$(x - x_s, x_k) = 0 \quad (k = 1, 2, \dots, m). \quad (17)$$

When we substitute in (17) for  $x_s$  its expression (16), we obtain:

[illegible]

Regarding this as a system of linear homogeneous equations with the non-zero solution  $c_1, c_2, \dots, c_m, -1$ , we equate the determinant of the system to zero and obtain (after transposition with respect to the main diagonal):<sup>10</sup>

$$\begin{vmatrix} (x_1 x_1) \dots (x_1 x_m) & x_1 \\ \dots & \dots \\ (x_m x_1) \dots (x_m x_m) & x_m \\ (x x_1) \dots (x x_m) & x_s \end{vmatrix} = 0. \quad (19)$$

When we separate from this determinant the term containing  $x_s$ , we obtain (in a readily understandable notation) :

$$x_s = \frac{\begin{array}{c} \begin{array}{c} x_1 \\ \vdots \\ x_m \\ 0 \end{array} \\ \begin{array}{c} (x x_1) \dots (x x_m) \\ \alpha \end{array} \end{array}}, \quad (20)$$

where  $G = G(x_1, x_2, \dots, x_m)$  is the Gramian of the vectors  $x_1, x_2, \dots, x_m$  (in virtue of the linear independence of these vectors,  $G \neq 0$ ). From (15) and (20), we find:

$$x_N = x - x_s = \frac{\begin{pmatrix} x x_1 \\ \vdots \\ x x_m \\ x \end{pmatrix}}{G}. \quad (21)$$

<sup>10</sup> The determinant on the left-hand side of (19) is a vector whose  $i$ -th coordinate is obtained by replacing all the vectors  $\mathbf{x}_1, \dots, \mathbf{x}_m, \mathbf{x}_g$  in the last column by their  $i$ -th coordinates ( $i=1, 2, \dots, n$ ); the coordinates are taken in an arbitrary basis. To justify the transition from (18) to (19), it is sufficient to replace the vectors  $\mathbf{x}_1, \dots, \mathbf{x}_m, \mathbf{x}_g$  by their  $i$ -th coordinates.

The formulas (20) and (21) express the projection  $x_s$  of  $x$  onto the subspace  $S$  and the projecting vector  $x_N$  in terms of the given vector  $x$  and the basis of  $S$ .

2. We draw attention to another important formula. We denote by  $h$  the length of the vector  $x_N$ . Then, by (15) and (21),

$$h^2 = (x_N x_N) = (x_N x) = \frac{\begin{vmatrix} G & (x_1 x) \\ & \vdots \\ & (x_m x) \\ (x x_1) \dots (x x_m) & (x x) \end{vmatrix}}{G},$$

i.e.,

$$h^2 = \frac{G(x_1, x_2, \dots, x_m, x)}{G(x_1, x_2, \dots, x_m)} \quad (22)$$

The quantity  $h$  can also be interpreted in the following way:

Let the vectors  $x_1, x_2, \dots, x_m, x$  issue from a single point and construct on these vectors as edges an  $(m+1)$ -dimensional parallelepiped. Then  $h$  is the height of this parallelepiped measured from the end of the edge  $x$  to the base  $S$  that passes through the edges  $x_1, x_2, \dots, x_m$ .

Let  $y$  be an arbitrary vector of  $S$  and  $x$  an arbitrary vector of  $R$ . If all vectors start from the origin of coordinates of an  $n$ -dimensional point space, then  $|x - y|$  and  $|x - x_s|$  are equal to the value of the slant height and the height respectively from the endpoint of  $x$  to the hyperplane  $S$ .<sup>11</sup> Therefore, when we set down that the height is shorter than the slant height, we have:<sup>12</sup>

$$h = |x - x_s| \leq |x - y|$$

(with equality only for  $y = x_s$ ). Thus, among all vectors  $y \in S$  the vector  $x_s$  deviates the least from the given vector  $x \in R$ . The quantity  $h = \sqrt{N(x - x_s)}$  is the mean-square error in the approximation  $x \approx x_s$ .<sup>13</sup>

## § 5. The Geometrical Meaning of the Gramian and Some Inequalities

1. We consider arbitrary vectors  $x_1, x_2, \dots, x_m$ . Let us assume, to begin with, that they are linearly independent. In this case the Gramian formed from any of these vectors is different from zero. Then, when we set, in accordance with (22),

<sup>11</sup> See the example on p. 248.

<sup>12</sup>  $N(x - y) = N(x_N + x_s - y) = N x_N + N(x_s - y) \geq N(x_N) = h^2$ .

<sup>13</sup> As regards the application of metrized functional spaces to problems of approximation of functions, see [1].

$$\frac{G(x_1, x_2, \dots, x_{p+1})}{G(x_1, x_2, \dots, x_p)} = h_p^2 > 0 \quad (p=1, 2, \dots, m-1), \quad (23)$$

and multiply these inequalities and the inequality

$$G(x_1) = (x_1 x_1) > 0, \quad (24)$$

we obtain

$$G(x_1, x_2, \dots, x_m) > 0.$$

Thus: *The Gramian of linearly independent vectors is positive; that of linearly dependent vectors is zero. Negative Gramians do not exist.*

Let us use the abbreviation  $G_p = G(x_1, x_2, \dots, x_p)$  ( $p=1, 2, \dots, m$ ). Then, from (23) and (24), we have

$$\sqrt{G_1} = |x_1| = V_1,$$

$$\sqrt{G_2} = V_1 h_1 = V_2,$$

where  $V_2$  is the area of the parallelogram spanned by  $x_1$  and  $x_2$ . Further,

$$\sqrt{G_3} = V_2 h_2 = V_3,$$

where  $V_3$  is the volume of the parallelepiped spanned by  $x_1, x_2, x_3$ . Continuing further, we find:

$$\sqrt{G_4} = V_3 h_3 = V_4$$

and, in general,

$$\sqrt{G_m} = V_{m-1} h_{m-1} = V_m. \quad (25)$$

It is natural to call  $V_m$  the volume of the  $m$ -dimensional parallelepiped spanned by the vectors  $x_1, x_2, \dots, x_m$ .<sup>14</sup>

We denote by  $x_{1k}, x_{2k}, \dots, x_{mk}$  the coordinates of  $x_k$  ( $k=1, 2, \dots, m$ ) in an orthonormal basis of  $R$  and set

$$X = \|x_{ik}\| \quad (i=1, 2, \dots, n; k=1, 2, \dots, m).$$

Then, in consequence of (10),

$$G_m = |X^T \bar{X}|$$

and therefore (see formula (25)),

$$V_m^2 = G_m = \sum_{1 \leq i_1 < i_2 < \dots < i_m \leq n} \text{mod} \begin{vmatrix} x_{i_1 1} & x_{i_1 2} & \dots & x_{i_1 m} \\ x_{i_2 1} & x_{i_2 2} & \dots & x_{i_2 m} \\ \dots & \dots & \dots & \dots \\ x_{i_m 1} & x_{i_m 2} & \dots & x_{i_m m} \end{vmatrix}^2. \quad (26)$$

<sup>14</sup> Formula (25) gives an inductive definition of the volume of an  $m$ -dimensional parallelepiped.

This equation has the following geometric meaning:

*The square of the volume of a parallelepiped is equal to the sum of the squares of the volumes of its projections on all the  $m$ -dimensional coordinate subspaces. In particular, for  $m = n$ , it follows from (26) that*

$$V_n^2 = \text{mod} \begin{vmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{vmatrix}. \quad (27)$$

The formulas (20), (21), (22), (26), and (27) solve a number of fundamental metrical problems of  $n$ -dimensional unitary and  $n$ -dimensional euclidean analytical geometry.

2. Let us return to the decomposition (15). This has the immediate consequence:

$$(\mathbf{x}\mathbf{x}) = (\mathbf{x}_S + \mathbf{x}_N, \mathbf{x}_S + \mathbf{x}_N) = (\mathbf{x}_S, \mathbf{x}_S) + (\mathbf{x}_N, \mathbf{x}_N) \geq (\mathbf{x}_N, \mathbf{x}_N) = h^2,$$

which, in conjunction with (22), gives an inequality (for arbitrary vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m, \mathbf{x}$ )

$$G(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m, \mathbf{x}) \leq G(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m) G(\mathbf{x}); \quad (28)$$

the equality sign holds if and only if  $\mathbf{x}$  is orthogonal to  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$ .

From this we easily obtain the so-called *Hadamard inequality*

$$G(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m) \leq G(\mathbf{x}_1) G(\mathbf{x}_2) \dots G(\mathbf{x}_m), \quad (29)$$

where the equality sign holds if and only if the vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$  are pairwise orthogonal. The inequality (29) expresses the following fact, which is geometrically obvious:

*The volume of a parallelepiped does not exceed the product of the lengths of its edges and is equal to it only when the parallelepiped is rectangular.*

Hadamard's inequality can be put into its usual form by setting  $m = n$  in (29) and introducing the determinant  $\Delta$  formed from the coordinates  $x_{1k}, x_{2k}, \dots, x_{nk}$  of the vectors  $\mathbf{x}_k$  ( $k = 1, 2, \dots, n$ ) in some orthonormal basis:

$$\Delta = \begin{vmatrix} x_{11} & \dots & x_{1n} \\ \dots & \dots & \dots \\ x_{n1} & \dots & x_{nn} \end{vmatrix}.$$

Then it follows from (27) and (29) that

$$|\Delta|^2 \leq \sum_{i=1}^n |x_{i1}|^2 \sum_{i=1}^n |x_{i2}|^2 \cdots \sum_{i=1}^n |x_{in}|^2. \quad (29')$$

3.<sup>15</sup> We now turn to the inequality

$$G(x_{1S}, x_{2S}, \dots, x_{mS}) \leq G(x_1, x_2, \dots, x_m) \quad (30)$$

If  $G(x_1, x_2, \dots, x_m) \neq 0$ , then the equality sign holds in (30) if and only if  $x_{iN} = 0$  ( $i = 1, 2, \dots, m$ ). If  $G(x_1, x_2, \dots, x_m) = 0$ , then (30) implies, of course, that  $G(x_{1S}, x_{2S}, \dots, x_{mS}) = 0$ .

In virtue of (25), the inequality (30) expresses the following geometric fact.

*The volume of the orthogonal projection of a parallelepiped onto a subspace  $S$  does not exceed the volume of the given parallelepiped; these volumes are equal if and only if the projecting parallelepiped lies in  $S$  or has zero volume.*

We prove (30) by induction on  $m$ .

The first step ( $m = 1$ ) is trivial and yields the inequality

$$G(x_{1S}) \leq G(x_1),$$

i.e.,  $|x_{1S}| \leq |x_1|$  (see Fig. 5 on page 248).

We write the volume  $\sqrt{G(x_1, x_2, \dots, x_m)}$  of our parallelepiped as the product of the 'base'  $\sqrt{G(x_1, x_2, \dots, x_{m-1})}$  by the distance  $h$  of the vertex of  $x_m$  from the base:

$$\sqrt{G(x_1, x_2, \dots, x_{m-1})} \cdot h = \sqrt{G(x_1, x_2, \dots, x_m)}. \quad (31)$$

If we now go over on the left-hand side of (31) from the vectors  $x_i$  to their projections  $x_{iS}$  ( $i = 1, 2, \dots, m$ ), then the first factor cannot increase, by the induction hypothesis, nor the second, by a simple geometric argument. But the product so obtained is the volume  $\sqrt{G(x_{1S}, x_{2S}, \dots, x_{mS})}$  of the parallelepiped projected onto the subspace  $S$ . Hence

$$\sqrt{G(x_{1S}, x_{2S}, \dots, x_{mS})} \leq \sqrt{G(x_1, x_2, \dots, x_m)}.$$

and by squaring both sides, we obtain (30).

Our condition for the equality sign to hold follows immediately from the proof.

<sup>15</sup> Subsections 3 and 4 have been modified in accordance with a correction published by the author in 1954 (Uspehi Mat. Nauk, vol. 9, no. 3).



4. Now we shall establish a generalization of Hadamard's inequality which comprises both the inequalities (28) and (29) :

$$G(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m) \leq G(\mathbf{x}_1, \dots, \mathbf{x}_p) G(\mathbf{x}_{p+1}, \dots, \mathbf{x}_m), \quad (32)$$

where the equality sign holds if and only if each vector  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$  is orthogonal to each of the vectors  $\mathbf{x}_{p+1}, \dots, \mathbf{x}_m$  or one of the determinants  $G(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p), G(\mathbf{x}_{p+1}, \dots, \mathbf{x}_m)$  vanishes.

The inequality (32) has the following geometric meaning :

*The volume of a parallelepiped does not exceed the product of the volumes of two complementary 'faces' and is equal to this product if and only if these faces are orthogonal or at least one of them has volume zero.*

Let us prove the inequality (32). Let  $p < m$ . If  $G(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p) = 0$ , then (32) holds with the equality sign. Let  $G(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p) \neq 0$ . Then the  $p$  vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$  are linearly independent and form a basis of a  $p$ -dimensional subspace  $T$  of  $R$ . The set of all vectors  $\mathbf{y}$  of  $R$  that are orthogonal to  $T$  are easily seen also to form a subspace of  $R$  (the so-called orthogonal complement of  $T$ ; for details, see § 8 of this Chapter). We denote it by  $S$ , and then  $R = T + S$ .

Since every vector of  $S$  is orthogonal to every vector of  $T$ , we can go over, in the Gramian  $G(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m)$ , whose square represents a certain volume, from the vectors  $\mathbf{x}_{p+1}, \dots, \mathbf{x}_m$  to their projections  $\mathbf{x}_{p+1S}, \dots, \mathbf{x}_{mS}$  onto the subspace  $S$ :

$$G(\mathbf{x}_1, \dots, \mathbf{x}_p, \mathbf{x}_{p+1}, \dots, \mathbf{x}_m) = G(\mathbf{x}_1, \dots, \mathbf{x}_p, \mathbf{x}_{p+1S}, \dots, \mathbf{x}_{mS}).$$

The same arguments show that the Gramian on the right-hand side of this equation can be split:

$$G(\mathbf{x}_1, \dots, \mathbf{x}_p, \mathbf{x}_{p+1S}, \dots, \mathbf{x}_{mS}) = G(\mathbf{x}_1, \dots, \mathbf{x}_p) G(\mathbf{x}_{p+1S}, \dots, \mathbf{x}_{mS}).$$

If we now go back from the projections to the original vectors and use (30), then we obtain

$$G(\mathbf{x}_1, \dots, \mathbf{x}_p) G(\mathbf{x}_{p+1S}, \dots, \mathbf{x}_{mS}) \leq G(\mathbf{x}_1, \dots, \mathbf{x}_p) G(\mathbf{x}_{p+1}, \dots, \mathbf{x}_m).$$

The equality sign holds in two cases: 1. When  $G(\mathbf{x}_{p+1}, \dots, \mathbf{x}_m) = 0$ , for then it is obvious that  $G(\mathbf{x}_{p+1S}, \dots, \mathbf{x}_{mS}) = 0$ ; and 2. When  $\mathbf{x}_{iS} = \mathbf{x}_i$  ( $i = 1, 2, 3, \dots, m$ ), i.e., when the vectors  $\mathbf{x}_{p+1}, \dots, \mathbf{x}_m$  belong to  $S$  or, what is the same, each vector  $\mathbf{x}_{p+1}, \dots, \mathbf{x}_m$  is orthogonal to every vector  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$  (the case  $G(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p) = 0$  has been considered at the beginning of the proof). By combining the last three relations we obtain the generalized Hadamard inequality (32) and the conditions for the equality sign to hold. This completes the proof.

5. The generalized Hadamard inequality (32) can also be put into analytic form.

Let  $\sum_{i,k=1}^n h_{ik} x_i \bar{x}_k$  be an arbitrary positive-definite hermitian form. By regarding  $x_1, x_2, \dots, x_n$  as the coordinates, in a basis  $e_1, e_2, \dots, e_n$ , of a vector  $\mathbf{x}$  in an  $n$ -dimensional space  $\mathbf{R}$ , we take  $\sum_{i,k=1}^n h_{ik} x_i \bar{x}_k$  as the fundamental metric form of  $\mathbf{R}$  (see p. 244). Then  $\mathbf{R}$  becomes a unitary space. We apply the generalized Hadamard inequality to the basis vectors  $e_1, e_2, \dots, e_n$ :

$$G(e_1, e_2, \dots, e_n) \leq G(e_1, \dots, e_p) G(e_{p+1}, \dots, e_n).$$

Setting  $H = \|h_{ik}\|_1^n$  and noting that  $(e_i, e_k) = h_{ik}$  ( $i, k = 1, 2, \dots, n$ ), we can rewrite the latter inequality as follows:

$$H \begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix} \leq H \begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} H \begin{pmatrix} p+1 & \dots & n \\ p+1 & \dots & n \end{pmatrix} \quad (p < n). \quad (33)$$

Here the equality sign holds if and only if  $h_{ik} = h_{ki} = 0$  ( $i = 1, 2, \dots, p$ ;  $k = p+1, \dots, n$ ).

The inequality (33) holds for the coefficient matrix  $H = \|h_{ik}\|_1^n$  of an arbitrary positive-definite hermitian form. In particular, (33) holds if  $H$  is the real coefficient matrix of a positive-definite quadratic form

$$\sum_{i,k=1}^n h_{ik} x_i x_k.^{16}$$

6. We remind the reader of Schwarz's inequality:†

For arbitrary vectors  $\mathbf{x}, \mathbf{y} \in \mathbf{R}$

$$|(\mathbf{x}\mathbf{y})|^2 \leq \mathbf{N}\mathbf{x}\mathbf{N}\mathbf{y}, \quad (34)$$

and the equality sign holds only if the vectors  $\mathbf{x}$  and  $\mathbf{y}$  differ only by a scalar factor

The validity of Schwarz's inequality follows easily from the inequality established above

$$G(\mathbf{x}, \mathbf{y}) = \begin{vmatrix} (\mathbf{x}\mathbf{x}) & (\mathbf{x}\mathbf{y}) \\ (\mathbf{y}\mathbf{x}) & (\mathbf{y}\mathbf{y}) \end{vmatrix} \geq 0.$$

By analogy with the scalar product of vectors in a three-dimensional euclidean space, we can introduce in an  $n$ -dimensional unitary space the

<sup>16</sup> An analytical approach to the generalized Hadamard inequality can be found in the book [17], § 8.

† In the Russian literature, this is known as Bunyakovskii's inequality.

'angle'  $\theta$  between the vectors  $\mathbf{x}$  and  $\mathbf{y}$  by defining<sup>17</sup>

$$\cos^2 \theta = \frac{|(\mathbf{x}\mathbf{y})|^2}{N\mathbf{x}N\mathbf{y}}.$$

From Schwarz's inequality it follows that  $\theta$  is real.

### § 6. Orthogonalization of a Sequence of Vectors

1. The smallest subspace containing the vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$  will be denoted by  $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p]$ . This subspace consists of all possible linear combinations  $c_1 \mathbf{x}_1 + c_2 \mathbf{x}_2 + \dots + c_p \mathbf{x}_p$  of the vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$  ( $c_1, c_2, c_3, \dots, c_p$  are complex numbers.)<sup>18</sup> If  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$  are linearly independent, then they form a basis of  $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p]$ . In that case, the subspace is  $p$ -dimensional.

Two sequences of vectors

$$\mathbf{X}: \mathbf{x}_1, \mathbf{x}_2, \dots,$$

$$\mathbf{Y}: \mathbf{y}_1, \mathbf{y}_2, \dots$$

containing an equal number of vectors, finite or infinite, will be called *equivalent* if for all  $p$

$$[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p] = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_p] \quad (p = 1, 2, \dots).$$

A sequence of vectors

$$\mathbf{X}: \mathbf{x}_1, \mathbf{x}_2, \dots$$

will be called *non-degenerate* if for every  $p$  the vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_p$  are linearly independent.

A sequence of vectors is called *orthogonal* if any two vectors of the sequence are orthogonal.

By *orthogonalization* of a sequence of vectors we mean a process of replacing the sequence by an equivalent orthogonal sequence.

**THEOREM 2:** *Every non-degenerate sequence of vectors can be orthogonalized. The orthogonalizing process leads to vectors that are uniquely determined to within scalar multiples.*

<sup>17</sup> In the case of a euclidean space, the angle  $\theta$  between the vectors  $\mathbf{x}$  and  $\mathbf{y}$  is defined by the formula

$$\cos \theta = \frac{(\mathbf{x}\mathbf{y})}{|\mathbf{x}||\mathbf{y}|}.$$

<sup>18</sup> In the case of a euclidean space, these numbers are real.

*Proof.* 1) Let us prove the second part of the theorem first. Suppose that two orthogonalizing sequences  $y_1, y_2, \dots (Y)$  and  $x_1, x_2, \dots (Z)$  are equivalent to one and the same non-degenerate sequence  $x_1, x_2, \dots (X)$ . Then  $Y$  and  $Z$  are equivalent to each other. Therefore for every  $p$  there exist numbers  $c_{p1}, c_{p2}, \dots, c_{pp}$  such that

$$x_p = c_{p1}y_1 + c_{p2}y_2 + \dots + c_{p,p-1}y_{p-1} + c_{pp}y_p \quad (p=1, 2, \dots).$$

When we form the scalar products of both sides of this equation by  $y_1, y_2, \dots, y_{p-1}$  and take account of the orthogonality of  $Y$  and of the relation

$$x_p \perp [x_1, x_2, \dots, x_{p-1}] \equiv [y_1, y_2, \dots, y_{p-1}],$$

we obtain  $c_{p1} = c_{p2} = \dots = c_{p,p-1} = 0$ , and therefore

$$x_p = c_{pp}y_p \quad (p=1, 2, \dots).$$

2) A concrete form of the orthogonalizing process for an arbitrary non-degenerate sequence of vectors  $x_1, x_2, \dots (X)$  is given by the following construction.

Let

$$S_p \equiv [x_1, x_2, \dots, x_p], \quad G_p = G(x_1, x_2, \dots, x_p) \quad (p=1, 2, \dots).$$

We project the vector  $x_p$  orthogonally onto the subspace  $S_{p-1}$  ( $p=1, 2, \dots$ ):<sup>10</sup>

$$x_p = x_{ps_{p-1}} + x_{pN}, \quad x_{ps_{p-1}} \in S_{p-1}, \quad x_{pN} \perp S_{p-1} \quad (p=1, 2, \dots).$$

We set

$$y_p = \lambda_p x_{pN} \quad (p=1, 2, \dots; x_{1N} = x_1),$$

where  $\lambda_p$  ( $p=1, 2, \dots$ ) are arbitrary non-zero numbers.

Then it is easily seen that

$$Y: y_1, y_2, \dots$$

is an orthogonal sequence equivalent to  $X$ . This proves Theorem 2.

By (21)

$$x_{pN} = \frac{\begin{vmatrix} & & x_1 \\ & & \vdots \\ G & & \vdots \\ & & x_{p-1} \end{vmatrix}}{G_{p-1}} \frac{x_p}{x_p} \quad (p=1, 2, \dots; G_0=1).$$

<sup>10</sup> For  $p=1$  we set  $x_{1S_0} = 0$  and  $x_{1N} = x_1$ .

Setting  $\lambda_p = G_{p-1}$  ( $p = 1, 2, \dots; G_0 = 1$ ), we obtain the following formulas for the vectors of the orthogonalized sequence:

$$y_1 = x_1, y_2 = \begin{vmatrix} (x_1 x_1) & x_1 \\ (x_2 x_1) & x_2 \end{vmatrix}, \dots, y_p = \begin{vmatrix} (x_1 x_1) & \dots & (x_1 x_{p-1}) & x_1 \\ \dots & \dots & \dots & \dots \\ (x_{p-1} x_1) & \dots & (x_{p-1} x_{p-1}) & x_{p-1} \\ (x_p x_1) & \dots & (x_p x_{p-1}) & x_p \end{vmatrix}, \dots \quad (35)$$

By (22),

$$N y_p = G_{p-1}^2 N x_{pN} = G_{p-1}^2 \cdot \frac{G_p}{G_{p-1}} = G_{p-1} G_p \quad (p = 1, 2, \dots; G_0 = 1). \quad (36)$$

Therefore, setting

$$z_p = \frac{y_p}{\sqrt{G_{p-1} G_p}} \quad (p = 1, 2, \dots), \quad (37)$$

we obtain an orthogonal sequence  $Z$  equivalent to the given sequence  $X$ .

*Example.* In the space of real functions that are sectionally continuous in the interval  $[-1, +1]$ , we define the scalar product

$$(f, g) = \int_{-1}^{+1} f(x) g(x) dx.$$

We consider the non-degenerate sequence of 'vectors'

$$1, x, x^2, x^3, \dots$$

We orthogonalize this sequence by the formulas (35):

$$y_0 \equiv 1, y_m = \begin{vmatrix} \frac{1}{1} & 0 & \frac{1}{3} & 0 & \frac{1}{5} & 0 & \dots & 1 \\ 0 & \frac{1}{3} & 0 & \frac{1}{5} & 0 & \frac{1}{7} & \dots & x \\ \frac{1}{3} & 0 & \frac{1}{5} & 0 & \frac{1}{7} & 0 & \dots & x^2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & x^m \end{vmatrix} \quad (m = 1, 2, \dots).$$

These orthogonal polynomials coincide, apart from constant factors, with the well-known Legendre polynomials:<sup>20</sup>

$$P_0(x) = 1, P_m(x) = \frac{1}{2^m m!} \frac{d^m (x^2 - 1)^m}{dx^m} \quad (m = 1, 2, \dots).$$

The same sequence of powers  $1, x, x^2, \dots$  in a different metric

<sup>20</sup> See [12], p. 77ff.

$$(f, g) = \int_a^b f(x)g(x)\tau(x)dx$$

(where  $\tau(x) \geq 0$  for  $a \leq x \leq b$ ) gives another sequence of orthogonal polynomials.

For example, if  $a = -1$ ,  $b = 1$  and  $\tau(x) = \frac{1}{\sqrt{1-x^2}}$ , then we obtain the Tchebyshev (Chebyshev) polynomials:

$$T_n(x) = \frac{1}{2^{n-1}} \cos(n \arccos x)$$

For  $a = -\infty$ ,  $b = +\infty$  and  $\tau(x) = e^{-x^2}$  we obtain the hermitian polynomials, etc.<sup>21</sup>

2. We shall now take note of the so-called Bessel inequality for an orthonormal sequence of vectors  $\mathbf{z}_1, \mathbf{z}_2, \dots (\mathbf{Z})$ . Let  $\mathbf{x}$  be an arbitrary vector. We denote by  $\xi_p$  the projection of  $\mathbf{x}$  onto  $\mathbf{z}_p$ :

$$\xi_p = (\mathbf{x}, \mathbf{z}_p) \quad (p = 1, 2, \dots).$$

Then the projection of  $\mathbf{x}$  onto the subspace  $S_p = [\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_p]$  can be represented in the form (see (20))

$$\mathbf{x}_{S_p} = \xi_1 \mathbf{z}_1 + \xi_2 \mathbf{z}_2 + \dots + \xi_p \mathbf{z}_p \quad (p = 1, 2, \dots).$$

But  $N\mathbf{x}_{S_p} = |\xi_1|^2 + |\xi_2|^2 + \dots + |\xi_p|^2 \leq N\mathbf{x}$ . Therefore, for every  $p$ ,

$$|\xi_1|^2 + |\xi_2|^2 + \dots + |\xi_p|^2 \leq N\mathbf{x}. \quad (38)$$

This is *Bessel's inequality*.

In the case of a space of finite dimension  $n$ , this inequality has a completely obvious geometrical meaning. For  $p = n$  it goes over into the theorem of Pythagoras

$$|\xi_1|^2 + |\xi_2|^2 + \dots + |\xi_n|^2 = |\mathbf{x}|^2.$$

In the case of an infinite-dimensional space and an infinite sequence  $\mathbf{Z}$ , it follows from (38) that the series  $\sum_{k=1}^{\infty} |\xi_k|^2$  converges and that

$$\sum_{k=1}^{\infty} |\xi_k|^2 \leq N\mathbf{x} = |\mathbf{x}|^2.$$

Let us form the series

<sup>21</sup> For further details see [12], Chapter II, § 9.

$$\sum_{k=1}^{\infty} \xi_k x_k.$$

For every  $p$  the  $p$ -th partial sum of this series,

$$\xi_1 x_1 + \xi_2 x_2 + \cdots + \xi_p x_p,$$

is the projection  $x_{S_p}$  of  $x$  onto the subspace

$$S_p = [x_1, x_2, \dots, x_p]$$

and is therefore the best approximation to the vector  $x$  in this subspace:

$$N(x - \sum_{k=1}^p \xi_k x_k) \leq N(x - \sum_{k=1}^p c_k x_k),$$

where  $c_1, c_2, \dots, c_p$  are arbitrary complex numbers. Let us calculate the corresponding mean-square-deviation  $\delta_p$ :

$$\delta_p^2 = N(x - \sum_{k=1}^p \xi_k x_k) = (x - \sum_{k=1}^p \xi_k x_k, x - \sum_{k=1}^p \xi_k x_k) = Nx - \sum_{k=1}^p |\xi_k|^2.$$

Hence

$$\lim_{p \rightarrow \infty} \delta_p^2 = Nx - \sum_{k=1}^{\infty} |\xi_k|^2.$$

If

$$\lim_{p \rightarrow \infty} \delta_p = 0,$$

then we say that the series  $\sum_{k=1}^{\infty} \xi_k x_k$  *converges in the mean* (or converges with respect to the norm) to the vector  $x$ .

In this case we have an equality for the vector  $x$  in  $R$  (the theorem of Pythagoras in an infinite-dimensional space!):

$$Nx = |x|^2 = \sum_{k=1}^{\infty} |\xi_k|^2. \quad (39)$$

If for every vector  $x$  of  $R$  the series  $\sum_{k=1}^{\infty} \xi_k x_k$  converges in the mean to  $x$ , then the orthonormal sequence of vectors  $x_1, x_2, \dots$  is called *complete*. In this case, when we replace  $x$  in (39) by  $x + y$  and use (39) three times, for  $N(x + y)$ ,  $Nx$ , and  $Ny$ , then we easily obtain:

$$(xy) = \sum_{k=1}^{\infty} \xi_k \bar{\eta}_k \quad [\xi_k = (xx_k), \eta_k = (yx_k); k = 1, 2, \dots]. \quad (40)$$

*Example.* We consider the space of all complex functions  $f(t)$  ( $t$  is a real variable) that are sectionally continuous in the closed interval  $[0, 2\pi]$ . Let us define the norm of  $f(t)$  by

$$Nf = \int_0^{2\pi} |f(t)|^2 dt.$$

Correspondingly, we have the formula

$$(f, g) = \int_0^{2\pi} f(t) \overline{g(t)} dt$$

for the scalar product of two functions  $f(t)$  and  $g(t)$ .

We take the infinite sequence of functions

$$\frac{1}{\sqrt{2\pi}} e^{ikt} \quad (k = 0, \pm 1, \pm 2, \dots).$$

These functions form an orthogonal sequence, because

$$\int_0^{2\pi} e^{i\mu t} \overline{e^{i\nu t}} dt = \int_0^{2\pi} e^{i(\mu-\nu)t} dt = \begin{cases} 0, & \text{for } \mu \neq \nu, \\ 2\pi, & \text{for } \mu = \nu. \end{cases}$$

The series

$$\sum_{k=-\infty}^{\infty} f_k e^{ikt}, \quad \left( f_k = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt; \quad (k = 0, \pm 1, \pm 2, \dots) \right)$$

converges in the mean to  $f(t)$  in the interval  $[0, 2\pi]$ . This series is called the *Fourier series* of  $f(t)$  and the coefficients  $f_k$  ( $k = 0, \pm 1, \pm 2, \dots$ ) are called the *Fourier coefficients* of  $f(t)$ .

In the theory of Fourier series it is proved that the system of functions  $e^{ikt}$  ( $k = 0, \pm 1, \pm 2, \dots$ ) is complete.<sup>22</sup>

The condition of completeness gives *Parseval's equality* (see (40))

$$\int_0^{2\pi} f(t) \overline{g(t)} dt = \sum_{k=-\infty}^{+\infty} \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt \int_0^{2\pi} \overline{g(t)} e^{ikt} dt.$$

If  $f(t)$  is a real function, then  $f_0$  is real, and  $f_k$  and  $f_{-k}$  are conjugate complex numbers. Setting

$$f_k = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-ikt} dt = \frac{1}{2} (a_k + ib_k),$$

<sup>22</sup> See, for example, [12], Chapter II.



where

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos kt \, dt, \quad b_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin kt \, dt \quad (k=0, 1, 2, \dots).$$

we have

$$f_k e^{ikt} + f_{-k} e^{-ikt} = a_k \cos kt + b_k \sin kt \quad (k=1, 2, \dots).$$

Therefore, for a real function  $f(t)$  the Fourier series assumes the form

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kt + b_k \sin kt) \quad \left( \begin{array}{l} a_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos kt \, dt, \\ b_k = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin kt \, dt, \end{array} \quad k=0, 1, 2, \dots \right).$$

## § 7. Orthonormal Bases

**1.** A basis of any finite-dimensional subspace  $S$  in a unitary or a euclidean space  $R$  is a non-degenerate sequence of vectors and therefore—by Theorem 2 of the preceding section—can be orthogonalized and normalized. Thus: *Every finite-dimensional subspace  $S$  (and, in particular, the whole space  $R$  if it is finite-dimensional) has an orthonormal basis.*

Let  $e_1, e_2, \dots, e_n$  be an orthonormal basis of  $R$ . We denote by  $x_1, x_2, x_3, \dots, x_n$  the coordinates of an arbitrary vector  $x$  in this basis:

$$x = \sum_{k=1}^n x_k e_k.$$

Multiplying both sides of this equation on the right by  $e_k$  and taking into account that the basis is orthonormal, we easily find:

$$x_k = (x e_k) \quad (k=1, 2, \dots, n);$$

i.e., *in an orthonormal basis the coordinates of a vector are equal to its projections onto the corresponding basis vectors:*

$$x = \sum_{k=1}^n (x e_k) e_k. \quad (41)$$

Let  $x_1, x_2, \dots, x_n$  and  $x'_1, x'_2, \dots, x'_n$  be the coordinates of one and the same vector  $x$  in two different orthonormal bases  $e_1, e_2, \dots, e_n$  and  $e'_1, e'_2, \dots, e'_n$  of a unitary space  $R$ . The formulas for the coordinate transformation have the form

$$x_i = \sum_{k=1}^n u_{ik} x'_k \quad (i = 1, 2, \dots, n). \quad (42)$$

Here the coefficients  $u_{1k}, u_{2k}, \dots, u_{nk}$  that form the  $k$ -th column of the matrix  $U = \|u_{ik}\|_1^n$  are easily seen to be the coordinates of the vector  $\mathbf{e}'_k$  in the basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ . Therefore, when we write down the condition for the basis  $\mathbf{e}'_1, \mathbf{e}'_2, \dots, \mathbf{e}'_n$  to be orthonormal in terms of coordinates (see (10)), we obtain the relations

$$\sum_{i=1}^n u_{ik} \bar{u}_{il} = \delta_{kl} = \begin{cases} 1, & \text{for } k=l, \\ 0, & \text{for } k \neq l. \end{cases} \quad (43)$$

A transformation (42) in which the coefficients satisfy the conditions (43) is called *unitary* and the corresponding matrix  $U$  is called a *unitary matrix*. Thus: *In an  $n$ -dimensional unitary space the transition from one orthonormal basis to another is effected by a unitary coordinate transformation.*

Let  $\mathbf{R}$  be an  $n$ -dimensional euclidean space. The transition from one orthonormal basis of  $\mathbf{R}$  to another is effected by a coordinate transformation

$$x_i = \sum_{k=1}^n v_{ik} x'_k \quad (i = 1, 2, \dots, n) \quad (44)$$

whose coefficients are connected by the relation

$$\sum_{i=1}^n v_{ik} v_{il} = \delta_{kl} \quad (k, l = 1, 2, \dots, n). \quad (45)$$

Such a coordinate transformation is called *orthogonal* and the corresponding matrix  $V$  is called an *orthogonal matrix*.

**2.** We note an interesting matrix method of writing the orthogonalizing process. Let  $A = \|a_{ik}\|_1^n$  be an arbitrary non-singular matrix ( $|A| \neq 0$ ) with complex elements. We consider a unitary space  $\mathbf{R}$  with an orthonormal basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  and define the linearly independent vectors  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$  by the equations

$$\mathbf{a}_k = \sum_{i=1}^n a_{ik} \mathbf{e}_i \quad (k = 1, 2, \dots, n).$$

Let us perform the orthogonalizing process on the vectors  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ . The orthonormal basis of  $\mathbf{R}$  so obtained we shall denote by  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ . Suppose we have

$$\mathbf{u}_k = \sum_{i=1}^n u_{ik} \mathbf{e}_i \quad (k = 1, 2, \dots, n).$$

Then

$$[a_1, a_2, \dots, a_p] = [u_1, u_2, \dots, u_p] \quad (p=1, 2, \dots, n),$$

i.e.,

$$\begin{aligned} a_1 &= c_{11}u_1, \\ a_2 &= c_{12}u_1 + c_{22}u_2, \\ &\dots \dots \dots \\ a_n &= c_{1n}u_1 + c_{2n}u_2 + \dots + c_{nn}u_n, \end{aligned}$$

where the  $c_{ik}$  ( $i, k=1, 2, \dots, n; i \leq k$ ) are certain complex numbers.

Setting  $c_{ik}=0$  for  $i > k$ , we have:

$$a_k = \sum_{p=1}^n c_{pk}u_p \quad (k=1, 2, \dots, n).$$

When we go over to coordinates and introduce the upper triangular matrix  $C = \|c_{ik}\|_1^n$  and the unitary matrix  $U = \|u_{ik}\|_1^n$ , we obtain

$$a_{ik} = \sum_{p=1}^n u_{ip}c_{pk} \quad (i, k=1, 2, \dots, n),$$

or

$$A = UC. \quad (*)$$

According to this formula: *Every non-singular matrix  $A = \|a_{ik}\|_1^n$  can be represented in the form of a product of a unitary matrix  $U$  and an upper triangular matrix  $C$ .*

Since the orthogonalizing process determines the vectors  $u_1, u_2, \dots, u_n$  uniquely, apart from scalar multipliers  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  ( $|\varepsilon_i|=1; i=1, 2, \dots, n$ ), the factors  $U$  and  $C$  in (\*) are uniquely determined apart from a diagonal factor  $M = \{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n\}$ :

$$U = U_1 M_1 \quad C = M^{-1} C_1.$$

This can also be shown directly.

*Note 1.* If  $A$  is a real matrix, the factors  $U$  and  $C$  in (\*) can be chosen to be real. In this case,  $U$  is an orthogonal matrix.

*Note 2.* The formula (\*) also remains valid for a singular matrix  $A$  ( $|A|=0$ ). This can be seen by setting  $A = \lim_{m \rightarrow \infty} A_m$ , where  $|A_m| \neq 0$  ( $m=1, 2, \dots$ ).

Then  $A_m = U_m C_m$  ( $m=1, 2, \dots$ ). When we select from the sequence  $\{U_m\}$  a convergent subsequence  $\{U_{m_p}\}$  ( $\lim_{p \rightarrow \infty} U_{m_p} = U$ ) and proceed to the limit, then we obtain from the equation  $A_{m_p} = U_{m_p} C_{m_p}$  for  $p \rightarrow \infty$  the required decomposition  $A = UC$ . However, in the case  $|A|=0$  the factors  $U$  and  $C$  are no longer uniquely determined to within a diagonal factor  $M$ .

Note 3. Instead of (\*) we can also obtain a formula

$$A = DW, \quad (**)$$

where  $D$  is a lower triangular matrix and  $W$  a unitary matrix. For when we apply the formula (\*) that was established above to the transposed matrix  $A^T$

$$A^T = UC$$

and then set  $W = U^T$ ,  $D = C^T$ , we obtain (\*\*).<sup>23</sup>

### § 8. The Adjoint Operator

1. Let  $A$  be a linear operator in an  $n$ -dimensional unitary space.

DEFINITION 4: A linear operator  $A^*$  is called adjoint to the operator  $A$  if and only if for any two vectors  $x, y$  of  $R$

$$(Ax, y) = (x, A^*y). \quad (46)$$

We shall show that for every linear operator  $A$  there exists one and only one adjoint operator  $A^*$ . To prove this, we take an orthonormal basis  $e_1, e_2, \dots, e_n$  in  $R$ . Then (see (41)) the required operator  $A^*$  and an arbitrary vector  $y$  of  $R$  must satisfy the equation

$$A^*y = \sum_{k=1}^n (A^*y, e_k) e_k.$$

By (46) this can be rewritten as follows:

$$A^*y = \sum_{k=1}^n (y, Ae_k) e_k. \quad (47)$$

We now take (47) as the definition of an operator  $A^*$ .

It is easy to verify that the operator  $A^*$  so defined is linear and satisfies (46) for arbitrary vectors  $x$  and  $y$  of  $R$ . Moreover, (47) determines the operator  $A^*$  uniquely. Thus the existence and uniqueness of the adjoint operator  $A^*$  is established.

Let  $A$  be a linear operator in a unitary space and let  $A = \|a_{ik}\|_1^n$  be the corresponding matrix in an orthonormal basis  $e_1, e_2, \dots, e_n$ . Then, by applying the formula (41) to the vector  $Ae_k = \sum_{i=1}^n a_{ik} e_i$ , we obtain

$$a_{ik} = (Ae_k, e_i) \quad (i, k = 1, 2, \dots, n). \quad (48)$$

<sup>23</sup> From the fact that  $U$  is unitary it follows that  $U^T$  is unitary, since the condition (43), written in matrix form  $U^T U = E$ , implies that  $U U^T = E$ .

Now let  $A^* = \|a_{ik}^*\|_1^n$  be the matrix corresponding to  $A^*$  in the same basis. Then, by (48),

$$a_{ik}^* = (A^* e_k, e_i) \quad (i, k = 1, 2, \dots, n). \quad (49)$$

From (48) and (49) it follows by (46) that

$$a_{ik}^* = \bar{a}_{ki} \quad (i, k = 1, 2, \dots, n),$$

i.e.,

$$A^* = \bar{A}^T.$$

The matrix  $A^*$  is the complex conjugate of the transpose of  $A$ . This matrix will be called the *adjoint* of  $A$ . (This is not to be confused with the adjoint of a matrix as defined on p. 14.)

Thus: *In an orthonormal basis adjoint matrices correspond to adjoint operators.*

The following properties of the adjoint operator follow from its definition:

1.  $(A^*)^* = A$ ,
2.  $(A + B)^* = A^* + B^*$ ,
3.  $(\alpha A)^* = \bar{\alpha} A^*$  ( $\alpha$  a scalar),
4.  $(AB)^* = B^* A^*$ .

2. We shall now introduce an important concept. Let  $S$  be an arbitrary subspace of  $R$ . We denote by  $T$  the set of all vectors  $y$  of  $R$  that are orthogonal to  $S$ . It is easy to see that  $T$  is a subspace of  $R$  and that every vector  $x$  of  $R$  can be represented uniquely in the form of a sum  $x = x_s + x_T$ , where  $x_s \in S$ ,  $x_T \in T$ , so that we have the resolution

$$R = S + T, \quad S \perp T.$$

We obtain this resolution by applying the decomposition (15) to the arbitrary vector  $x$  of  $R$ .  $T$  is called the *orthogonal complement* of  $S$ . Obviously,  $S$  is the orthogonal complement of  $T$ . We write  $S \perp T$ , meaning by this that each vector of  $S$  is orthogonal to every vector of  $T$ .

Now we can formulate the fundamental property of the adjoint operator:

5. *If a subspace  $S$  is invariant with respect to  $A$ , then the orthogonal complement  $T$  of the subspace is invariant with respect to  $A^*$ .*

For let  $\mathbf{x} \in S$ ,  $\mathbf{y} \in T$ . Then it follows from  $A\mathbf{x} \in S$  that  $(A\mathbf{x}, \mathbf{y}) = 0$  and hence by (46) that  $(\mathbf{x}, A^*\mathbf{y}) = 0$ . Since  $\mathbf{x}$  is an arbitrary vector of  $S$ ,  $A^*\mathbf{y} \in T$ , and this is what we had to prove.

We introduce the following definition:

DEFINITION 5: Two systems of vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m$  and  $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m$  are called bi-orthogonal if

$$(\mathbf{x}_i, \mathbf{y}_k) = \delta_{ik} \quad (i, k = 1, 2, \dots, m), \quad (50)$$

where  $\delta_{ik}$  is the Kronecker symbol.

Now we shall prove the following proposition:

6. If  $A$  is a linear operator of simple structure, then the adjoint operator  $A^*$  is also of simple structure, and complete systems of characteristic vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  and  $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$  of  $A$  and  $A^*$  can be chosen such that they are bi-orthogonal:

$$A\mathbf{x}_i = \lambda_i \mathbf{x}_i, \quad A^*\mathbf{y}_i = \mu_i \mathbf{y}_i, \quad (\mathbf{x}_i, \mathbf{y}_k) = \delta_{ik} \quad (i, k = 1, 2, \dots, n).$$

For let  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  be a complete system of characteristic vectors of  $A$ . We use the notation

$$S_k = [\mathbf{x}_1, \dots, \mathbf{x}_{k-1}, \mathbf{x}_{k+1}, \dots, \mathbf{x}_n] \quad (k = 1, 2, \dots, n).$$

Consider the one-dimensional orthogonal complement  $T_k = [\mathbf{y}_k]$  to the  $(n-1)$ -dimensional subspace  $S_k$  ( $k = 1, 2, \dots, n$ ). Then  $T_k$  is invariant with respect to  $A^*$ :

$$A^*\mathbf{y}_k = \mu_k \mathbf{y}_k, \quad \mathbf{y}_k \neq \mathbf{0} \quad (k = 1, 2, \dots, n).$$

From  $S_k \perp \mathbf{y}_k$  it follows that  $(\mathbf{x}_k, \mathbf{y}_k) \neq 0$ , because otherwise the vector  $\mathbf{y}_k$  would have to be the null vector. Multiplying  $\mathbf{x}_k, \mathbf{y}_k$  ( $k = 1, 2, \dots, n$ ) by suitable numerical factors we obtain

$$(\mathbf{x}_i, \mathbf{y}_k) = \delta_{ik} \quad (i, k = 1, 2, \dots, n).$$

From the bi-orthogonality of the systems  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  and  $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$  it follows that the vectors of each system are linearly independent.

We mention one further proposition:

7. If the operators  $A$  and  $A^*$  have a common characteristic vector, then the corresponding characteristic values are complex conjugates.

For let  $A\mathbf{x} = \lambda\mathbf{x}$  and  $A^*\mathbf{x} = \mu\mathbf{x}$  ( $\mathbf{x} \neq \mathbf{0}$ ). Then, setting  $\mathbf{y} = \mathbf{x}$  in (46), we have  $\lambda(\mathbf{x}, \mathbf{x}) = \bar{\mu}(\mathbf{x}, \mathbf{x})$  and hence  $\lambda = \bar{\mu}$ .

## § 9. Normal Operators in a Unitary Space

1. DEFINITION 6. A linear operator  $A$  is called *normal* if it commutes with its adjoint:

$$AA^* = A^*A. \quad (51)$$

DEFINITION 7. A linear operator  $H$  is called *hermitian* if it is equal to its adjoint:

$$H^* = H. \quad (52)$$

DEFINITION 8. A linear operator  $U$  is called *unitary* if it is inverse to its adjoint:

$$UU^* = E \quad (53)$$

Note that a unitary operator can be regarded as an isometric operator in a hermitian space, i.e., as an operator preserving the metric.

For suppose that for arbitrary vectors  $x$  and  $y$  of  $R$

$$(Ux, Uy) = (x, y). \quad (54)$$

Then by (46)

$$(U^*Ux, y) = (x, y)$$

and therefore, since  $y$  is arbitrary,

$$U^*Ux = x,$$

i.e.,  $U^*U = E$ , or  $U^* = U^{-1}$ . Conversely, (53) implies (54).

From (53) and (54) it follows that 1. the product of two unitary operators is itself a unitary operator, 2. the unit operator  $E$  is unitary, and 3. the inverse of a unitary operator is also unitary. Therefore the set of all unitary operators is a group.<sup>24</sup> This is called the *unitary group*.

Hermitian operators and unitary operators are special cases of a normal operator.

2. We have

THEOREM 3: Every linear operator  $A$  can be represented in the form

$$A = H_1 + iH_2, \quad (55)$$

where  $H_1$  and  $H_2$  are hermitian operators (the 'hermitian components' of  $A$ ). The hermitian components are uniquely determined by  $A$ . The operator  $A$  is normal if and only if its hermitian components  $H_1$  and  $H_2$  are permutable.

<sup>24</sup> See footnote 13 on p. 18.

*Proof.* Suppose that (55) holds. Then

$$A^* = H_1 - iH_2. \quad (56)$$

From (55) and (56) we have:

$$H_1 = \frac{1}{2}(A + A^*), \quad H_2 = \frac{1}{2i}(A - A^*). \quad (57)$$

Conversely, the formulas (57) define hermitian operators  $H_1$  and  $H_2$  connected with  $A$  by (55).

Now let  $A$  be a normal operator:  $AA^* = A^*A$ . Then it follows from (57) that  $H_1H_2 = H_2H_1$ . Conversely, from  $H_1H_2 = H_2H_1$  it follows by (55) and (56) that  $AA^* = A^*A$ . This completes the proof.

The representation of an arbitrary linear operator  $A$  in the form (55) is an analogue to the representation of a complex number  $z$  in the form  $x_1 + ix_2$ , where  $x_1$  and  $x_2$  are real.

Suppose that in some orthonormal basis the operators  $A$ ,  $H$ , and  $U$  correspond to the matrices  $A$ ,  $H$ , and  $U$ . Then the operator equations

$$AA^* = A^*A, \quad H^* = H, \quad UU^* = E \quad (58)$$

correspond to the matrix equations

$$AA^* = A^*A, \quad H^* = H, \quad UU^* = E. \quad (59)$$

Therefore we define a matrix as *normal* if it commutes with its adjoint, as *hermitian* if it is equal to its adjoint, and finally as *unitary* if it is inverse to its adjoint.

Then: *In an orthonormal basis a normal (hermitian, unitary) operator corresponds to a normal (hermitian, unitary) matrix.*

A hermitian matrix  $H = \|h_{ik}\|_1^n$  is, by (59), characterized by the following relation among its elements:

$$h_{ki} = \overline{h_{ik}} \quad (i, k = 1, 2, \dots, n);$$

i.e., a hermitian matrix is always the coefficient matrix of some hermitian form (see § 1).

A unitary matrix  $U = \|u_{ik}\|_1^n$  is, by (59), characterized by the following relations among its elements:

$$\sum_{j=1}^n u_{ij} \overline{u_{kj}} = \delta_{ik} \quad (i, k = 1, 2, \dots, n). \quad (60)$$



Since  $UU^* = E$  implies that  $U^*U = E$ , from (60) there follow the equivalent relations:

$$\sum_{j=1}^n u_{jk} \bar{u}_{ji} = \delta_{ik} \quad (i, k = 1, 2, \dots, n). \quad (61)$$

Equation (60) expresses the 'orthonormality' of the rows and equation (61) that of the columns of the matrix  $U = \| u_{ik} \|_1^{25}$

A unitary matrix is the coefficient matrix of some unitary transformation (see § 7).

### § 10. The Spectra of Normal, Hermitian, and Unitary Operators

1. As a preliminary, we establish a property of permutable operators in the form of a lemma.

**LEMMA 1:** *Permutable operators  $A$  and  $B$  ( $AB = BA$ ) always have a common characteristic vector.*

*Proof.* Let  $x$  be a characteristic vector of  $A$ :  $Ax = \lambda x$ ,  $x \neq 0$ . Then, since  $A$  and  $B$  are permutable,

$$AB^k x = \lambda B^k x \quad (k = 0, 1, 2, \dots). \quad (62)$$

Suppose that in the sequence of vectors

$$x, Bx, B^2x, \dots$$

the first  $p$  are linearly independent, while the  $(p+1)$ -th vector  $B^p x$  is a linear combination of the preceding ones. Then  $S \equiv [x, Bx, \dots, B^{p-1}x]$  is a subspace invariant with respect to  $B$ , so that in this subspace  $S$  there exists a characteristic vector  $y$  of  $B$ :  $By = \mu y$ ,  $y \neq 0$ . On the other hand, (62) shows that the vectors  $x, Bx, \dots, B^{p-1}x$  are characteristic vectors of  $A$  corresponding to one and the same characteristic value  $\lambda$ . Therefore every linear combination of these vectors, and in particular  $y$ , is a characteristic vector of  $A$  corresponding to  $\lambda$ . Thus we have proved the existence of a common characteristic vector of the operators  $A$  and  $B$ .

Let  $A$  be an arbitrary normal operator in an  $n$ -dimensional hermitian space  $R$ . In that case  $A$  and  $A^*$  are permutable and therefore have a common characteristic vector  $x_1$ . Then (see § 8, 7.)

<sup>25</sup> Thus, orthonormality of the columns of the matrix  $U$  is a consequence of the orthonormality of the rows, and vice versa.

$$Ax_1 = \lambda_1 x_1, \quad A^* x_1 = \bar{\lambda}_1 x_1 \quad (x_1 \neq 0).$$

We denote by  $S_1$  the one-dimensional subspace containing the vector  $x_1$  ( $S_1 = [x_1]$ ) and by  $T_1$  the orthogonal complement of  $S_1$  in  $R$ :

$$R = S_1 + T_1, \quad S_1 \perp T_1.$$

Since  $S_1$  is invariant with respect to  $A$  and  $A^*$ ,  $T_1$  is also invariant with respect to these operators (see § 8, 5.). Therefore, by Lemma 1, the permutable operators  $A$  and  $A^*$  have a common characteristic vector  $x_2$  in  $T_1$ :

$$Ax_2 = \lambda_2 x_2, \quad A^* x_2 = \bar{\lambda}_2 x_2 \quad (x_2 \neq 0).$$

Obviously,  $x_1 \perp x_2$ . Setting  $S_2 = [x_1, x_2]$  and

$$R = S_2 + T_2, \quad S_2 \perp T_2,$$

we establish in a similar way the existence of a common characteristic vector  $x_3$  of  $A$  and  $A^*$  in  $T_2$ . Obviously  $x_1 \perp x_3$  and  $x_2 \perp x_3$ . Continuing this process, we obtain  $n$  pairwise orthogonal common characteristic vectors  $x_1, x_2, \dots, x_n$  of  $A$  and  $A^*$ :

$$\left. \begin{aligned} Ax_k &= \lambda_k x_k, \quad A^* x_k = \bar{\lambda}_k x_k \quad (x_k \neq 0), \\ (x_i, x_k) &= 0, \quad \text{for } i \neq k \end{aligned} \right\} \quad (i, k = 1, 2, \dots, n). \quad (63)$$

The vectors  $x_1, x_2, \dots, x_n$  can be normalized without violating (63).

Thus we have proved that a normal operator always has a complete orthonormal system of characteristic vectors.<sup>26</sup>

Since  $\lambda_k = \lambda_l$  always implies that  $\bar{\lambda}_k = \bar{\lambda}_l$ , it follows from (63) that:

1. If  $A$  is a normal operator, every characteristic vector of  $A$  is a characteristic vector of the adjoint operator  $A^*$ , i.e., if  $A$  is a normal operator then  $A$  and  $A^*$  have the same characteristic vectors.

Suppose now, conversely, that a linear operator  $A$  has a complete orthonormal system of characteristic vectors:

$$Ax_k = \lambda_k x_k, \quad (x_i, x_k) = \delta_{ik} \quad (i, k = 1, 2, \dots, n).$$

We shall show that  $A$  is then a normal operator. For let us set:

$$y_i = Ax_i - \bar{\lambda}_i x_i.$$

Then

$$\begin{aligned} (x_k, y_l) &= (x_k, Ax_l - \bar{\lambda}_l x_l) = (Ax_k, x_l) - \bar{\lambda}_l (x_k, x_l) \\ &= (\lambda_k - \bar{\lambda}_l) \delta_{kl} = 0 \quad (k, l = 1, 2, \dots, n). \end{aligned}$$

Hence it follows that

<sup>26</sup> Here, and in what follows, we mean by a complete orthonormal system of vectors an orthonormal system of  $n$  vectors, where  $n$  is the dimension of the space.

$$y_l = A^* x_l - \bar{\lambda}_l x_l = 0 \quad (l=1, 2, \dots, n),$$

i.e., that (63) holds.

But then

$$AA^* x_k = \lambda_k \bar{\lambda}_k x_k \text{ and } A^* A x_k = \lambda_k \bar{\lambda}_k x_k \quad (k=1, 2, \dots, n),$$

or

$$AA^* = A^* A.$$

Thus we have obtained the following 'internal' (spectral) characterization of a normal operator  $A$  (apart from the 'external' one:  $AA^* = A^* A$ ):

**THEOREM 4:** *A linear operator is normal if and only if it has a complete orthonormal system of characteristic values.*

In particular, we have shown that a normal operator is always of simple structure.

Let  $A$  be a normal operator with the characteristic values  $\lambda_1, \lambda_2, \dots, \lambda_n$ . Using the Lagrange interpolation formula, we define two polynomials  $p(\lambda)$  and  $q(\lambda)$  by the conditions

$$p(\lambda_k) = \bar{\lambda}_k, \quad q(\bar{\lambda}_k) = \lambda_k \quad (k=1, 2, \dots, n).$$

Then by (63)

$$A^* = p(A), \quad A = q(A^*); \quad (64)$$

i.e.:

2. *If  $A$  is a normal operator, then each of the operators  $A$  and  $A^*$  can be represented as a polynomial in the other; these two polynomials are determined by the characteristic values of  $A$ .*

Let  $S$  be an invariant subspace of  $R$  for a normal operator  $A$  and let  $R = S + T$ ,  $S \perp T$ . Then by § 8, 5., the subspace  $T$  is invariant with respect to  $A^*$ . But  $A = q(A^*)$ , where  $q(\lambda)$  is a polynomial. Therefore  $T$  is also invariant with respect to  $A$ . Thus:

3. *If  $S$  is an invariant subspace with respect to a normal operator  $A$  and  $T$  is the orthogonal complement of  $S$ , then  $T$  is also an invariant subspace for  $A$ .*

2. Let us now discuss the spectrum of a hermitian operator. Since a hermitian operator  $H$  is a special form of a normal operator, by what we have proved it has a complete orthonormal system of characteristic vectors:

$$Hx_k = \lambda_k x_k, \quad (x_k, x_l) = \delta_{kl} \quad (k, l=1, 2, \dots, n). \quad (65)$$

From  $H^* = H$  it follows that

$$\bar{\lambda}_k = \lambda_k \quad (k=1, 2, \dots, n), \quad (66)$$

i.e., all the characteristic values of a hermitian operator  $\mathbf{H}$  are real.

It is not difficult to see that, conversely, a normal operator with real characteristic values is always hermitian. For from (65), (66), and

$$\mathbf{H}^* \mathbf{x}_k = \lambda_k \mathbf{x}_k \quad (k=1, 2, \dots, n)$$

it follows that

$$\mathbf{H}^* \mathbf{x}_k = \mathbf{H} \mathbf{x}_k \quad (k=1, 2, \dots, n),$$

i.e.

$$\mathbf{H}^* = \mathbf{H}.$$

We have obtained the following 'internal' characterization of a hermitian operator (apart from the 'external' one:  $\mathbf{H}^* = \mathbf{H}$ ):

**THEOREM 5:** *A linear operator  $\mathbf{H}$  is hermitian if and only if it has a complete orthonormal system of characteristic vectors with real characteristic values.*

Let us now discuss the spectrum of a unitary operator. Since a unitary operator  $\mathbf{U}$  is normal, it has a complete orthonormal system of characteristic vectors

$$\mathbf{U} \mathbf{x}_k = \lambda_k \mathbf{x}_k, \quad (\mathbf{x}_k, \mathbf{x}_l) = \delta_{kl} \quad (k, l=1, 2, \dots, n), \quad (67)$$

where

$$\mathbf{U}^* \mathbf{x}_k = \bar{\lambda}_k \mathbf{x}_k \quad (k=1, 2, \dots, n). \quad (68)$$

From  $\mathbf{U} \mathbf{U}^* = \mathbf{E}$  we find:

$$\lambda_k \bar{\lambda}_k = 1. \quad (69)$$

Conversely, from (67), (68), and (69) it follows that  $\mathbf{U} \mathbf{U}^* = \mathbf{E}$ . Thus, among the normal operators a unitary operator is distinguished by the fact that all its characteristic values have modulus 1.

We have thus obtained the following 'internal' characterization of a unitary operator (apart from the 'external' one:  $\mathbf{U} \mathbf{U}^* = \mathbf{E}$ ):

**THEOREM 6:** *A linear operator is unitary if and only if it has a complete orthonormal system of characteristic vectors with characteristic values of modulus 1.*

Since in an orthonormal basis a normal (hermitian, unitary) matrix corresponds to a normal (hermitian, unitary) operator, we obtain the following propositions:

**THEOREM 4':** *A matrix  $\mathbf{A}$  is normal if and only if it is unitarily similar to a diagonal matrix:*

$$\mathbf{A} = \mathbf{U} \|\lambda_i \delta_{ik}\|_1 \mathbf{U}^{-1} \quad (\mathbf{U}^* = \mathbf{U}^{-1}). \quad (70)$$

**THEOREM 5':** *A matrix  $H$  is hermitian if and only if it is unitarily similar to a diagonal matrix with real diagonal elements:*

$$H = U \|\lambda_i \delta_{ik}\|_1^n U^{-1} \quad (U^* = U^{-1}; \lambda_i = \bar{\lambda}_i; i = 1, 2, \dots, n). \quad (71)$$

**THEOREM 6':** *A matrix  $U$  is unitary if and only if it is unitarily similar to a diagonal matrix with diagonal elements of modulus 1:*

$$U = U_1 \|\lambda_i \delta_{ik}\|_1^n U_1^{-1} \quad (U_1^* = U_1^{-1}; |\lambda_i| = 1; i = 1, 2, \dots, n). \quad (72)$$

## § 11. Positive-Semidefinite and Positive-Definite Hermitian Operators

1. We introduce the following definition:

**DEFINITION 9:** *A hermitian operator  $H$  is called positive semidefinite if for every vector  $\mathbf{x}$  of  $R$*

$$(H\mathbf{x}, \mathbf{x}) \geq 0,$$

*and positive definite if for every vector  $\mathbf{x} \neq \mathbf{o}$  of  $R$*

$$(H\mathbf{x}, \mathbf{x}) > 0.$$

If a vector  $\mathbf{x}$  is given by its coordinates  $x_1, x_2, \dots, x_n$  in an arbitrary orthonormal basis, then  $(H\mathbf{x}, \mathbf{x})$ , as is easy to see, is a hermitian form in the variables  $x_1, x_2, \dots, x_n$ ; and to a positive-semidefinite (positive-definite) operator there corresponds a positive-semidefinite (positive-definite) hermitian form (see § 1).

We choose an orthonormal basis  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  of characteristic vectors of  $H$ :

$$H\mathbf{x}_k = \lambda_k \mathbf{x}_k, \quad (\mathbf{x}_k, \mathbf{x}_l) = \delta_{kl} \quad (k, l = 1, 2, \dots, n). \quad (73)$$

Then, setting  $\mathbf{x} = \sum_{k=1}^n \xi_k \mathbf{x}_k$ , we have

$$(H\mathbf{x}, \mathbf{x}) = \sum_{k=1}^n \lambda_k |\xi_k|^2 \quad (k = 1, 2, \dots, n).$$

Hence we easily deduce the 'internal' characterizations of positive-semidefinite and positive-definite operators:

**THEOREM 7:** *A hermitian operator is positive semidefinite (positive definite) if and only if all its characteristic values are non-negative (positive).*

From what we have shown, it follows that a positive-definite hermitian operator is non-singular and positive semidefinite.

Let  $H$  be a positive-semidefinite hermitian operator. The equation (73) holds for  $H$  with  $\lambda_k \geq 0$  ( $k = 1, 2, \dots, n$ ). We set  $\varrho_k = \sqrt{\lambda_k} \geq 0$  ( $k = 1, 2, 3, \dots, n$ ) and define a linear operator  $F$  by the equation

$$F x_k = \varrho_k x_k \quad (k = 1, 2, \dots, n). \quad (74)$$

Then  $F$  is also a positive-semidefinite operator and

$$F^2 = H. \quad (75)$$

We shall call the positive-semidefinite hermitian operator  $F$  connected with  $H$  by (75) the *arithmetical square root of  $H$*  and shall denote it by

$$F = \sqrt{H}.$$

If  $H$  is positive definite, then  $F$  is also positive definite.

We define the Lagrange interpolation polynomial  $g(\lambda)$  by the equations

$$g(\lambda_k) = \varrho_k (= \sqrt{\lambda_k}) \quad (k = 1, 2, \dots, n) \quad (76)$$

Then from (73), (74), and (76) it follows that:

$$F = g(H). \quad (77)$$

The latter equation shows that  $\sqrt{H}$  is a polynomial in  $H$  and is uniquely determined when the positive-semidefinite hermitian operator  $H$  is given (the coefficients of  $g(\lambda)$  depend on the characteristic values of  $H$ ).

2. Examples of positive-semidefinite hermitian operators are  $AA^*$  and  $A^*A$ , where  $A$  is an arbitrary linear operator in the given space. Indeed, for an arbitrary vector  $x$ ,

$$\begin{aligned} (AA^*x, x) &= (A^*x, A^*x) \geq 0, \\ (A^*Ax, x) &= (Ax, Ax) \geq 0. \end{aligned}$$

If  $A$  is non-singular, then  $AA^*$  and  $A^*A$  are positive-definite hermitian operators.

The operators  $AA^*$  and  $A^*A$  are sometimes called the *left norm* and *right norm* of  $A$ .  $\sqrt{AA^*}$  and  $\sqrt{A^*A}$  are called the *left modulus* and *right modulus* of  $A$ .

For a normal operator the left and right norms, and hence the left and right moduli, are equal.<sup>27</sup>

<sup>27</sup> For a detailed study of normal operators, see [168]. In this paper necessary and sufficient conditions for the product of two normal operators to be normal are established.

## § 12. Polar Decomposition of a Linear Operator in a Unitary Space. Cayley's Formulas

1. We shall prove the following theorem:<sup>28</sup>

**THEOREM 8:** *Every linear operator  $A$  in a unitary space can be represented in the forms*

$$A = HU, \quad (78)$$

$$A = U_1 H_1, \quad (79)$$

where  $H, H_1$  are positive-semidefinite hermitian operators and  $U, U_1$  are unitary operators.  $A$  is normal if and only if in (78) (or (79)) the factors  $H$  and  $U$  ( $H_1$  and  $U_1$ ) are permutable.

*Proof.* From (78) and (79) it follows that  $H$  and  $H_1$  are the left and right moduli, respectively, of  $A$ .

For

$$AA^* = HUU^*H = H^2, \quad A^*A = H_1U_1^*U_1H_1 = H_1^2.$$

Note that it is sufficient to establish (78), since by applying this decomposition to  $A^*$  we obtain  $A^* = HU$  and hence

$$A = U^{-1}H.$$

i.e., the decomposition (79) for  $A$ .

We begin by establishing (78) in the special case where  $A$  is non-singular. ( $|A| \neq 0$ ). We set:

$$H = \sqrt{AA^*} \text{ (here } |H|^2 = |A|^2 \neq 0), \quad U = H^{-1}A$$

and verify that  $U$  is unitary.

$$UU^* = H^{-1}AA^*H^{-1} = H^{-1}H^2H^{-1} = E.$$

Note that in this case not only the first factor  $H$  in (78), but also the second factor  $U$  is uniquely determined by the non-singular operator  $A$ .

We now consider the general case where  $A$  may be singular.

First of all we observe that a complete orthonormal system of characteristic vectors of the right norm of  $A$  is always transformed by  $A$  into an orthogonal system of vectors. For let

$$A^*Ax_k = \varrho_k^2 x_k \quad [(x_k, x_l) = \delta_{kl}, \varrho_k \geq 0; k, l = 1, 2, \dots, n].$$

Then

$$(Ax_k, Ax_l) = (A^*Ax_k, x_l) = \varrho_k^2 (x_k, x_l) = 0 \quad (k \neq l).$$

<sup>28</sup> See [168], p. 77.

Here

$$|Ax_k|^2 = (Ax_k, Ax_k) = \varrho_k^2 \quad (k=1, 2, \dots, n).$$

Therefore there exists an orthonormal system of vectors  $x_1, x_2, \dots, x_n$  such that

$$Ax_k = \varrho_k x_k \quad [(x_k, x_l) = \delta_{kl}; k, l=1, 2, \dots, n]. \quad (80)$$

We define linear operators  $H$  and  $U$  by the equations

$$Ux_k = x_k, \quad Hx_k = \varrho_k x_k.$$

From (80) and (81) we find:

$$A = HU.$$

Here  $H$  is, by (81), a positive-semidefinite hermitian operator, because it has a complete orthonormal system of characteristic vectors  $x_1, x_2, \dots, x_n$  with non-negative characteristic values  $\varrho_1, \varrho_2, \dots, \varrho_n$ ; and  $U$  is a unitary operator, because it carries the orthonormal system of vectors  $x_1, x_2, \dots, x_n$  into the orthonormal system  $x_1, x_2, \dots, x_n$ .

Thus we can take it as proved that an arbitrary linear operator  $A$  has decompositions (78) and (79), that the hermitian factors  $H$  and  $H_1$  are always uniquely determined by  $A$  (they are the left and right moduli of  $A$ , respectively) and that the unitary factors  $U$  and  $U_1$  are uniquely determined only when  $A$  is non-singular.

From (78) we find easily:

$$AA^* = H^2, \quad A^*A = U^{-1}H^2U. \quad (82)$$

If  $A$  is a normal operator ( $AA^* = A^*A$ ), then it follows from (82) that

$$H^2U = UH^2. \quad (83)$$

Since  $H = \sqrt{H^2} = g(H^2)$  (see § 11), (83) shows that  $U$  and  $H$  commute. Conversely, if  $H$  and  $U$  commute, then it follows from (82) that  $A$  is normal. This completes the proof of the theorem.<sup>29</sup>

<sup>29</sup> If the characteristic values  $\lambda_1, \lambda_2, \dots, \lambda_n$  and  $\varrho_1, \varrho_2, \dots, \varrho_n$  of the linear operator  $A$  and its left modulus  $H = \sqrt{AA^*}$  (by (82)  $\varrho_1, \varrho_2, \dots, \varrho_n$  are also the characteristic values of the right modulus  $H_1 = \sqrt{A^*A}$ ) are so numbered that

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|, \quad \varrho_1 \geq \varrho_2 \geq \dots \geq \varrho_n.$$

then (see [379], or [153] and [296]) the following inequality of Weyl holds:

$$|\lambda_1| \leq \varrho_1, \quad |\lambda_1| + |\lambda_2| \leq \varrho_1 + \varrho_2, \quad \dots, \quad |\lambda_1| + \dots + |\lambda_n| \leq \varrho_1 + \dots + \varrho_n.$$



It is hardly necessary to mention that together with the operator equations (78) and (79) the corresponding matrix equations hold.

The decompositions (78) and (79) are analogues to the representation of a complex number  $z$  in the form  $z = ru$ , where  $r = |z|$  and  $|u| = 1$ .

2. Now let  $x_1, x_2, \dots, x_n$  be a complete orthonormal system of characteristic vectors of the arbitrary unitary operator  $U$ . Then

$$Ux_k = e^{if_k} x_k, \quad (x_k, x_l) = \delta_{kl} \quad (k, l = 1, 2, \dots, n). \quad (84)$$

where the  $f_k$  ( $k = 1, 2, \dots, n$ ) are real numbers. We define a hermitian operator  $F$  by the equations

$$Fx_k = f_k x_k \quad (k = 1, 2, \dots, n). \quad (85)$$

From (84) and (85) it follows that:<sup>30</sup>

$$U = e^{iF}. \quad (86)$$

Thus, a unitary operator  $U$  is always representable in the form (86), where  $F$  is a hermitian operator. Conversely, if  $F$  is a hermitian operator, then  $U = e^{iF}$  is unitary.

The decompositions (78) and (79) together with (86) give the following equations:

$$A = He^{iF}, \quad (87)$$

$$A = e^{iF} H_1 \quad (88)$$

where  $H, F, H_1$ , and  $F_1$  are hermitian operators, with  $H$  and  $H_1$  positive semi-definite.

The decompositions (87) and (88) are analogues to the representation of a complex number  $z$  in the form  $z = re^{i\varphi}$ , where  $r \geq 0$  and  $\varphi$  are real numbers.

*Note.* In (86), the operator  $F$  is not uniquely determined by  $U$ . For  $F$  is defined by means of the numbers  $f_k$  ( $k = 1, 2, \dots, n$ ) and we can add to each of these numbers an arbitrary multiple of  $2\pi$  without changing the original equations (84). By choosing these multiples of  $2\pi$  suitably we can assume that  $e^{if_k} = e^{if_l}$  always implies that  $f_k = f_l$  ( $1 \leq k, l \leq n$ ). Then we can determine the interpolation polynomial  $g(\lambda)$  by the equations

$$g(e^{if_k}) = f_k \quad (k = 1, 2, \dots, n). \quad (89)$$

<sup>30</sup>  $e^{iF} = r(F)$ , where  $r(\lambda)$  is the Lagrange interpolation polynomial for the function  $e^{if_k}$  at the places  $f_1, f_2, \dots, f_n$ .

From (84), (85), and (89) it follows that

$$\mathbf{F} = g(\mathbf{U}) = g(e^{i\mathbf{F}}). \quad (90)$$

Similarly we can normalize the choice of  $\mathbf{F}_1$  so that

$$\mathbf{F}_1 = h(\mathbf{U}_1) = h(e^{i\mathbf{F}_1}), \quad (91)$$

where  $h(\lambda)$  is a polynomial.

By (90) and (91), the permutability of  $\mathbf{H}$  and  $\mathbf{U}$  ( $\mathbf{H}_1$  and  $\mathbf{U}_1$ ) implies that of  $\mathbf{H}$  and  $\mathbf{F}$  ( $\mathbf{H}_1$  and  $\mathbf{F}_1$ ), and vice versa. Therefore, by Theorem 8,  $\mathbf{A}$  is normal if and only if in (87)  $\mathbf{H}$  and  $\mathbf{F}$  (or, in (88),  $\mathbf{H}_1$  and  $\mathbf{F}_1$ ) are permutable, provided the characteristic values of  $\mathbf{F}$  (or  $\mathbf{F}_1$ ) are suitably normalized.

The formula (86) is based on the fact that the functional dependence

$$\mu = e^{if} \quad (92)$$

carries  $n$  arbitrary numbers  $f_1, f_2, \dots, f_n$  on the real axis into certain numbers  $\mu_1, \mu_2, \dots, \mu_n$  on the unit circle  $|\mu| = 1$ , and vice versa.

The transcendental dependence (92) can be replaced by the rational dependence

$$\mu = \frac{1 + if}{1 - if}, \quad (93)$$

which carries the real axis  $f = \bar{f}$  into the circle  $|\mu| = 1$ ; here the point at infinity on the real axis goes over into the point  $\mu = -1$ . From (93), we find:

$$f = i \frac{1 - \mu}{1 + \mu}. \quad (94)$$

Repeating the arguments which have led us to the formula (86), we obtain from (93) and (94) the pair of inverse formulas:

$$\left. \begin{aligned} \mathbf{U} &= (\mathbf{E} + i\mathbf{F})(\mathbf{E} - i\mathbf{F})^{-1}, \\ \mathbf{F} &= i(\mathbf{E} - \mathbf{U})(\mathbf{E} + \mathbf{U})^{-1}, \end{aligned} \right\} \quad (95)$$

We have thus obtained *Cayley's formulas*. These formulas establish a one-to-one correspondence between arbitrary hermitian operators  $\mathbf{F}$  and those unitary operators  $\mathbf{U}$  that do not have the characteristic value  $-1$ .<sup>31</sup>

<sup>31</sup> The exceptional value  $-1$  can be replaced by any number  $\mu_0$  ( $|\mu_0| = 1$ ). For this purpose, we have to take instead of (93) a fractional-linear function mapping the real axis  $f = \bar{f}$  onto the circle  $|\mu| = 1$  and carrying the point  $f = \infty$  into  $\mu = \mu_0$ . The formulas (94) and (95) can be modified correspondingly.

The formulas (86), (87), (88), and (95) are obviously valid when we replace all the operators by the corresponding matrices.

### § 13. Linear Operators in a Euclidean Space

1. We consider an  $n$ -dimensional euclidean space  $\mathbf{R}$ . Let  $\mathbf{A}$  be a linear operator in  $\mathbf{R}$ .

DEFINITION 10: The linear operator  $\mathbf{A}^\tau$  is called the *transposed operator* of  $\mathbf{A}$  (or the *transpose* of  $\mathbf{A}$ ) if for any two vectors  $\mathbf{x}$  and  $\mathbf{y}$  of  $\mathbf{R}$ :

$$(\mathbf{A}\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \mathbf{A}^\tau\mathbf{y}). \quad (96)$$

The existence and uniqueness of the transposed operator is established in exactly the same way as was done in § 8 for the adjoint operator in a unitary space.

The transposed operator has the following properties:

1.  $(\mathbf{A}^\tau)^\tau = \mathbf{A}$ ,
2.  $(\mathbf{A} + \mathbf{B})^\tau = \mathbf{A}^\tau + \mathbf{B}^\tau$ ,
3.  $(\alpha\mathbf{A})^\tau = \alpha\mathbf{A}^\tau$  ( $\alpha$  a real number),
4.  $(\mathbf{AB})^\tau = \mathbf{B}^\tau\mathbf{A}^\tau$ .

We introduce a number of definitions.

DEFINITION 11: A linear operator  $\mathbf{A}$  is called *normal* if

$$\mathbf{AA}^\tau = \mathbf{A}^\tau\mathbf{A}.$$

DEFINITION 12: A linear operator  $\mathbf{S}$  is called *symmetric* if

$$\mathbf{S}^\tau = \mathbf{S}.$$

DEFINITION 13: A symmetric operator  $\mathbf{S}$  is called *positive semidefinite* if for every vector  $\mathbf{x}$  of  $\mathbf{R}$

$$(\mathbf{S}\mathbf{x}, \mathbf{x}) \geq 0.$$

DEFINITION 14: A symmetric operator  $\mathbf{S}$  is called *positive definite* if for every vector  $\mathbf{x} \neq \mathbf{o}$  of  $\mathbf{R}$

$$(\mathbf{S}\mathbf{x}, \mathbf{x}) > 0.$$

DEFINITION 15: A linear operator  $\mathbf{K}$  is called *skew-symmetric* if

$$\mathbf{K}^\tau = -\mathbf{K}.$$

An arbitrary linear operator  $A$  can always be represented uniquely in the form

$$A = S + K, \quad (97)$$

where  $S$  is symmetric and  $K$  is skew-symmetric.

For it follows from (97) that

$$A^T = S - K. \quad (98)$$

From (97) and (98) we have:

$$S = \frac{1}{2} (A + A^T), \quad K = \frac{1}{2} (A - A^T). \quad (99)$$

Conversely, (99) defines a symmetric operator  $S$  and a skew-symmetric operator  $K$  for which (97) holds.

$S$  and  $K$  are called respectively the *symmetric component* and the *skew-symmetric component* of  $A$ .

DEFINITION 16: An operator  $Q$  is called *orthogonal* if it preserves the metric of the space, i.e., if for any two vectors  $x, y$  of  $R$

$$(Qx, Qy) = (x, y). \quad (100)$$

By (96), equation (100) can be written as:  $(x, Q^T Q y) = (x, y)$ . Hence

$$Q^T Q = E. \quad (101)$$

Conversely, (101) implies (100) (for arbitrary vectors  $x, y$ ).<sup>32</sup> From (101) it follows that:  $|Q|^2 = 1$ , i.e.,

$$|Q| = \pm 1.$$

We shall call  $Q$  an *orthogonal operator of the first kind* (or *proper*) if  $|Q| = 1$  and *of the second kind* (or *improper*) if  $|Q| = -1$ .

Symmetric, skew-symmetric, and orthogonal operators are special forms of a normal operator.

We consider an arbitrary orthonormal basis in the given euclidean space. Suppose that in this basis  $A$  corresponds to the matrix  $A = \|a_{ik}\|_1^n$  (here all the  $a_{ik}$  are real numbers). The reader will have no difficulty in showing that the transposed operator  $A^T$  corresponds in this basis to the transposed matrix  $A^T = \|a_{ik}^T\|_1^n$ , where  $a_{ik}^T = a_{ki}$  ( $i, k = 1, 2, \dots, n$ ). Hence it follows that in an orthonormal basis a normal operator  $A$  corresponds to a normal

<sup>32</sup> The orthogonal operators in a euclidean space form a group, the so-called orthogonal group.

matrix  $A$  ( $AA^T = A^T A$ ), a symmetric operator  $S$  to a symmetric matrix  $S = \| s_{ik} \|_1^n$  ( $S^T = S$ ), a skew-symmetric operator  $K$  to a skew-symmetric matrix  $K = \| k_{ij} \|_1^n$  ( $K^T = -K$ ) and, finally, an orthogonal operator  $Q$  to an orthogonal matrix  $Q = \| q_{ik} \|_1^n$  ( $QQ^T = E$ ).<sup>33</sup>

Just as was done in § 8 for the adjoint operator, we can here make the following statement for the transposed operator:

*If a subspace  $S$  of  $R$  is invariant with respect to a linear operator  $A$ , then the orthogonal complement  $T$  of  $S$  in  $R$  is invariant with respect to  $A^T$ .*

**2.** For the study of linear operators in a euclidean space  $R$ , we extend  $R$  to a unitary space  $\tilde{R}$ . This extension is made in the following way:

1. The vectors of  $R$  are called 'real' vectors.

2. We introduce 'complex' vectors  $z = x + iy$ , where  $x$  and  $y$  are real, i.e.,  $x \in R, y \in R$ .

3. The operations of addition of complex vectors and of multiplication by a complex number are defined in the natural way. Then the set of all complex vectors forms an  $n$ -dimensional vector space  $\tilde{R}$  over the field of complex numbers which contains  $R$  as a subspace.

4. In  $\tilde{R}$  we introduce a hermitian metric such that in  $R$  it coincides with the existing euclidean metric. The reader can easily verify that the required hermitian metric is given in the following way:

If  $z = x + iy, w = u + iv$  ( $x, y, u, v \in R$ ), then

$$(zw) = (xu) + (yv) + i[(yu) - (xv)].$$

Setting  $\bar{z} = x - iy$  and  $\bar{w} = u - iv$ , we have

$$(\bar{z}\bar{w}) = \overline{(zw)}.$$

If we choose a real basis, i.e., a basis of  $R$ , then  $\tilde{R}$  will be the set of all vectors with complex coordinates and  $R$  the set of all vectors with real coordinates in this basis.

Every linear operator  $A$  in  $R$  extends uniquely to a linear operator in  $\tilde{R}$ :

$$A(x + iy) = Ax + iAy.$$

**3.** Among all the linear operators of  $\tilde{R}$  those that are obtainable as the result of such an extension of operators of  $R$  can be characterized by the fact that they carry  $R$  into  $R$  ( $AR \subset R$ ). These operators are called *real*.

<sup>33</sup> The papers [138], [262a], [170b] are devoted to the study of the structure of orthogonal matrices. Orthogonal matrices, like orthogonal operators, are called proper and improper according as  $|Q| = +1$  or  $|Q| = -1$ .

In a real basis real operators are determined by real matrices, i.e., matrices with real elements.

A real operator  $A$  carries conjugate complex vectors  $z = x + iy$ ,  $\bar{z} = x - iy$  ( $x, y \in \mathbf{R}$ ) into conjugate complex vectors:

$$Az = Ax + iAy, A\bar{z} = Ax - iAy \quad (Ax, Ay \in \mathbf{R})$$

The secular equation of a real operator has real coefficients, so that when it has a root  $\lambda$  of multiplicity  $p$  it also has the root  $\bar{\lambda}$  with the multiplicity  $p$ . From  $Az = \lambda z$  it follows that  $A\bar{z} = \bar{\lambda}\bar{z}$ , i.e., to conjugate characteristic values there correspond conjugate characteristic vectors.

The two-dimensional space  $[z, \bar{z}]$  has a real basis:

$$x = \frac{1}{2}(z + \bar{z}), \quad y = \frac{1}{2i}(z - \bar{z}).$$

We shall call the plane in  $\mathbf{R}$  spanned by this basis an *invariant plane* of  $A$  corresponding to the pair of characteristic values  $\lambda, \bar{\lambda}$ .

Let  $\lambda = \mu + iv$ . Then it is easy to see that

$$Ax = \mu x - vy,$$

$$Ay = vx + \mu y.$$

We consider a real operator  $A$  of simple structure with the characteristic values:

$$\lambda_{2k-1} = \mu_k + iv_k, \lambda_{2k} = \mu_k - iv_k, \lambda_l = \mu_l \quad (k = 1, 2, \dots, q; l = 2q + 1, \dots, n),$$

where  $\mu_k, v_k, \mu_l$  are real and  $v_k \neq 0$  ( $k = 1, 2, \dots, q$ ).

Then the characteristic vectors  $z_1, z_2, \dots, z_n$  corresponding to these characteristic values can be chosen such that

$$z_{2k-1} = x_k + iy_k, \quad z_{2k} = x_k - iy_k, \quad z_l = x_l \quad (102)$$

$$(k = 1, 2, \dots, q; l = 2q + 1, \dots, n).$$

The vectors

$$x_1, y_1, x_2, y_2, \dots, x_q, y_q, x_{2q+1}, \dots, x_n \quad (103)$$

form a basis of the euclidean space  $\mathbf{R}$ . Here

<sup>34</sup> If to the characteristic value  $\lambda$  of the real operator  $A$  there correspond the linearly independent characteristic vectors  $z_1, z_2, \dots, z_p$ , then to the characteristic value  $\bar{\lambda}$  there correspond the linearly independent characteristic vectors  $\bar{z}_1, \bar{z}_2, \dots, \bar{z}_p$ .

$$\begin{aligned} Ax_k &= \mu_k x_k - \nu_k y_k, & (k=1, 2, \dots, q) \\ Ay_k &= \nu_k x_k + \mu_k y_k, & (l=2q+1, \dots, n) \\ Ax_l &= \mu_l x_l \end{aligned} \quad (104)$$

In the basis (103) there corresponds to the operator  $A$  the real quasi-diagonal matrix

$$\left\{ \left\| \begin{array}{cc} \mu_1 & \nu_1 \\ -\nu_1 & \mu_1 \end{array} \right\|, \dots, \left\| \begin{array}{cc} \mu_q & \nu_q \\ -\nu_q & \mu_q \end{array} \right\|, \mu_{2q+1}, \dots, \mu_n \right\}. \quad (105)$$

Thus: *For every operator  $A$  of simple structure in a euclidean space there exists a basis in which  $A$  corresponds to a matrix of the form (105). Hence it follows that: A real matrix of simple structure is real-similar to a canonical matrix of the form (105):*

$$A = T \left\{ \left\| \begin{array}{cc} \mu_1 & \nu_1 \\ -\nu_1 & \mu_1 \end{array} \right\|, \dots, \left\| \begin{array}{cc} \mu_q & \nu_q \\ -\nu_q & \mu_q \end{array} \right\|, \mu_{2q+1}, \dots, \mu_n \right\} T^{-1} \quad (T = \bar{T}). \quad (106)$$

The transposed operator  $A^T$  of  $A$  in  $\mathbf{R}$  upon extension becomes the adjoint operator  $A^*$  of  $A$  in  $\tilde{\mathbf{R}}$ . Therefore: *Normal, symmetric, skew-symmetric, and orthogonal operators in  $\mathbf{R}$  after the extension become normal, hermitian, hermitian multiplied by  $i$ , and unitary real operators in  $\tilde{\mathbf{R}}$ .*

It is easy to show that for a normal operator  $A$  in a euclidean space a canonical basis can be chosen as an orthonormal basis (103) for which (104) holds.<sup>35</sup> Therefore a real normal matrix is always real-similar and orthogonally-similar to a matrix of the form (105):

$$A = Q \left\{ \left\| \begin{array}{cc} \mu_1 & \nu_1 \\ -\nu_1 & \mu_1 \end{array} \right\|, \dots, \left\| \begin{array}{cc} \mu_q & \nu_q \\ -\nu_q & \mu_q \end{array} \right\|, \mu_{2q+1}, \dots, \mu_n \right\} Q^{-1} \quad (107)$$

$$(Q = Q^T = \bar{Q}).$$

All the characteristic values of a symmetric operator  $S$  in a euclidean space are real, since after the extension the operator becomes hermitian. For a symmetric operator  $S$  we must set  $q=0$  in (104). Then we obtain:

$$Sx_l = \mu_l x_l \quad [(x_l, x_l) = \delta_{kl}; k, l = 1, 2, \dots, n]. \quad (108)$$

*A symmetric operator  $S$  in a euclidean space always has an orthonormal system of characteristic vectors with real characteristic values.*<sup>36</sup> Therefore:

<sup>35</sup> The orthonormality of the basis (102) in the hermitian metric implies the orthonormality of the basis (103) in the corresponding euclidean metric.

<sup>36</sup> The symmetric operator  $S$  is positive semidefinite if in (108) all  $\mu_l \geq 0$  and positive definite if all  $\mu_l > 0$ .

A real symmetric matrix is always real-similar and orthogonally-similar to a diagonal matrix:

$$S = Q \{\mu_1, \mu_2, \dots, \mu_n\} Q^{-1} \quad (Q = Q^{\tau-1} = \bar{Q}). \quad (109)$$

All the characteristic values of a skew-symmetric operator  $K$  in a euclidean space are pure imaginary (after the extension the operator is  $i$  times a hermitian operator). For a skew-symmetric operator we must set in (104);

$$\mu_1 = \mu_2 = \dots = \mu_q = \mu_{2q+1} = \dots = \mu_n = 0$$

then the formulas assume the form

$$\begin{aligned} Kx_k &= -y_k, \\ Ky_k &= x_k, \quad (k=1, 2, \dots, q; \quad l=2q+1, \dots, n), \\ Kx_l &= 0 \end{aligned} \quad (110)$$

Since  $K$  is a normal operator, the basis (103) can be assumed to be orthonormal. Thus: *Every real skew-symmetric matrix is real-similar and orthogonally-similar to a canonical skew-symmetric matrix:*

$$K = Q \left\{ \left\| \begin{array}{cc} 0 & y_1 \\ -y_1 & 0 \end{array} \right\|, \dots, \left\| \begin{array}{cc} 0 & y_q \\ -y_q & 0 \end{array} \right\|, 0, \dots, 0 \right\} Q^{-1} \quad (Q = Q^{\tau-1} = \bar{Q}). \quad (111)$$

All the characteristic values of an orthogonal operator  $Q$  in a euclidean space are of modulus 1 (upon extension the operator becomes unitary). Therefore in the case of an orthogonal operator we must set in (104):

$$\mu_k^2 + \nu_k^2 = 1, \quad \mu_l = \pm 1 \quad (k=1, 2, \dots, q; \quad l=2q+1, \dots, n).$$

For this basis (103) can be assumed to be orthonormal. The formulas (104) can be represented in the form

$$\begin{aligned} Qx_k &= x_k \cos \varphi_k - y_k \sin \varphi_k, \\ Qy_k &= x_k \sin \varphi_k + y_k \cos \varphi_k, \\ Qx_l &= \pm x_l, \end{aligned} \quad \left( \begin{array}{l} k=1, 2, \dots, q, \\ l=2q+1, \dots, n \end{array} \right) \quad (112)$$

From what we have shown, it follows that: *Every real orthogonal matrix is real-similar and orthogonally-similar to a canonical orthogonal matrix:*

$$Q = Q_1 \left\{ \left\| \begin{array}{cc} \cos \varphi_1 & \sin \varphi_1 \\ -\sin \varphi_1 & \cos \varphi_1 \end{array} \right\|, \dots, \left\| \begin{array}{cc} \cos \varphi_q & \sin \varphi_q \\ -\sin \varphi_q & \cos \varphi_q \end{array} \right\|, \pm 1, \dots, \pm 1 \right\} Q_1^{-1} \quad (113)$$

$$(Q_1 = Q_1^{\tau-1} = \bar{Q}_1)$$



*Example.* We consider an arbitrary finite rotation around the point  $O$  in a three-dimensional space. It carries a directed segment  $\overrightarrow{OA}$  into a directed segment  $\overrightarrow{OB}$  and can therefore be regarded as an operator  $Q$  in a three-dimensional vector space (formed by all possible segments  $\overrightarrow{OA}$ ). This operator is linear and orthogonal. Its determinant is  $+1$ , since  $Q$  does not change the orientation of the space.

Thus,  $Q$  is a proper orthogonal operator. For this operator the formulas (112) look as follows:

$$Qx_1 = x_1 \cos \varphi - y_1 \sin \varphi,$$

$$Qy_1 = x_1 \sin \varphi + y_1 \cos \varphi,$$

$$Qx_2 = \pm x_2.$$

From the equation  $|Q| = 1$  it follows that  $Qx_2 = x_2$ . This means that all the points on the line through  $O$  in the direction of  $x_2$  remain fixed. Thus we have obtained the *Theorem of Euler-D'Alembert*:

*Every finite rotation of a rigid body around a fixed point can be obtained as a finite rotation by an angle  $\varphi$  around some fixed axis passing through that point.*

#### § 14. Polar Decomposition of an Operator and the Cayley Formulas in a Euclidean Space

1. In § 12 we established the polar decomposition of a linear operator in a unitary space. In exactly the same way we obtain the polar decomposition of a linear operator in a euclidean space.

**THEOREM 9.** *Every linear operator  $A$  is representable in the form of a product*<sup>37</sup>

$$A = SQ \quad (114)$$

$$A = Q_1 S_1 \quad (115)$$

where  $S, S_1$  are positive-semidefinite symmetric and  $Q, Q_1$  are orthogonal operators; here  $S = \sqrt{AA^T} = g(AA^T)$ ,  $S_1 = \sqrt{A^T A} = h(A^T A)$ , where  $g(\lambda)$  and  $h(\lambda)$  are real polynomials.

$A$  is a normal operator if and only if  $S$  and  $Q$  ( $S_1$  and  $Q_1$ ) are permutable.

Similar statements hold for matrices.

<sup>37</sup> As in Theorem 8, the operators  $S$  and  $S_1$  are uniquely determined by  $A$ . If  $A$  is non-singular, then the orthogonal factors  $Q$  and  $Q_1$  are also uniquely determined.

Let us point out the geometrical content of the formulas (114) and (115). We let the vectors of an  $n$ -dimensional euclidean point space issue from the origin of the coordinate system. Then every vector is the radius vector of some point of the space. The orthogonal transformation realized by the operator  $Q$  (or  $Q_1$ ) is a 'rotation' in this space, because it preserves the euclidean metric and leaves the origin of the coordinate system fixed.<sup>38</sup> The symmetric operator  $S$  (or  $S_1$ ) represents a 'dilatation' of the  $n$ -dimensional space (i.e., a 'stretching' along  $n$  mutually perpendicular directions with stretching factors  $\varrho_1, \varrho_2, \dots, \varrho_n$  that are, in general, distinct ( $\varrho_1, \varrho_2, \dots, \varrho_n$  are arbitrary non-negative numbers)). According to the formulas (114) and (115), every linear homogeneous transformation of an  $n$ -dimensional euclidean space can be obtained by carrying out in succession some rotation and some dilatation (in any order).

**2.** Just as was done in the preceding section for a unitary operator, we now consider some representations of an orthogonal operator in a euclidean space  $R$ .

Let  $K$  be an arbitrary skew-symmetric operator ( $K^T = -K$ ) and let

$$Q = e^K. \quad (116)$$

Then  $Q$  is a proper orthogonal operator. For

$$Q^T = e^{K^T} = e^{-K} = Q^{-1}$$

and

$$|Q| = 1.^{39}$$

Let us show that every proper orthogonal operator is representable in the form (116). For this purpose we take the corresponding orthogonal matrix  $Q$ . Since  $|Q| = 1$ , we have, by (113),<sup>40</sup>

<sup>38</sup> For  $|Q| = 1$  this is a proper rotation; but for  $|Q| = -1$  it is a combination of a rotation and a reflection in a coordinate plane.

<sup>39</sup> If  $k_1, k_2, \dots, k_n$  are the characteristic values of  $K$ , then  $\mu_1 = e^{k_1}, \mu_2 = e^{k_2}, \dots, \mu_n = e^{k_n}$  are the characteristic values of  $Q = e^K$ ; moreover

$$|Q| = \mu_1 \mu_2 \dots \mu_n = e^{\sum_{i=1}^n k_i} = 1,$$

since

$$\sum_{i=1}^n k_i = 0.$$

<sup>40</sup> Among the characteristic values of a proper orthogonal matrix  $Q$  there is an even number equal to  $-1$ . The diagonal matrix  $\begin{vmatrix} -1 & 0 \\ 0 & -1 \end{vmatrix}$  can be written in the form  $\begin{vmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{vmatrix}$  for  $\varphi = \pi$ .

$$Q = Q_1 \left\{ \left\| \begin{array}{cc} \cos \varphi_1 & \sin \varphi_1 \\ -\sin \varphi_1 & \cos \varphi_1 \end{array} \right\|, \dots, \left\| \begin{array}{cc} \cos \varphi_p & \sin \varphi_q \\ -\sin \varphi_p & \cos \varphi_q \end{array} \right\|, +1, \dots, +1 \right\} Q_1^{-1} \quad (117)$$

$$(Q_1 = (Q_1^T)^{-1} = \bar{Q}_1).$$

We define the skew-symmetric matrix  $K$  by the equation

$$K = Q_1 \left\{ \left\| \begin{array}{cc} 0 & \varphi_1 \\ -\varphi_1 & 0 \end{array} \right\|, \dots, \left\| \begin{array}{cc} 0 & \varphi_q \\ -\varphi_q & 0 \end{array} \right\|, 0, \dots, 0 \right\} Q_1^{-1}. \quad (118)$$

Since

$$\exp \left\{ \left\| \begin{array}{cc} 0 & \varphi \\ -\varphi & 0 \end{array} \right\| \right\} = \left\| \begin{array}{cc} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{array} \right\|,$$

it follows from (117) and (118) that

$$Q = e^K. \quad (119)$$

The matrix equation (119) implies the operator equation (116).

In order to represent an improper orthogonal operator we introduce a special operator  $W$  which is defined in an orthonormal basis  $e_1, e_2, \dots, e_n$  by the equations

$$We_1 = e_1, \dots, We_{n-1} = e_{n-1}, We_n = -e_n. \quad (120)$$

$W$  is an improper orthogonal operator. If  $Q$  is an arbitrary improper orthogonal operator then  $W^{-1}Q$  and  $QW^{-1}$  are proper and therefore representable in the form  $e^K$  and  $e^{K_1}$ , where  $K$  and  $K_1$  are skew-symmetric operators. Hence we obtain the formulas for an improper orthogonal operator

$$Q = We^K = e^{K_1} W. \quad (121)$$

The basis  $e_1, e_2, \dots, e_n$  in (120) can be chosen such that it coincides with the basis  $x_k, y_k, x_l$  ( $k = 1, 2, \dots, q; l = 2q + 1, \dots, n$ ) in (110) and (112). The operator  $W$  so defined is permutable with  $K$ ; therefore the two formulas (121) merge into one

$$Q = We^K \quad (W = W^T = W^{-1}; \quad K^T = -K, WK = KW). \quad (122)$$

Let us now turn to the Cayley formulas, which establish a connection between orthogonal and skew-symmetric operators in a euclidean space. The formula

$$Q = (E - K)(E + K)^{-1}, \quad (123)$$

as is easily verified, carries the skew-symmetric operator  $K$  into the orthogonal operator  $Q$ . (123) enables us to express  $K$  in terms of  $Q$ :

$$K = (E - Q)(E + Q)^{-1}. \quad (124)$$

The formulas (123) and (124) establish a one-to-one correspondence between the skew-symmetric operators and those orthogonal operators that do not have the characteristic value  $-1$ . Instead of (123) and (124) we can take the formulas

$$Q = -(E - K)(E + K)^{-1}, \quad (125)$$

$$K = (E + Q)(E - Q)^{-1}. \quad (126)$$

In this case the number  $+1$  plays the role of the exceptional value.

**3.** The polar decomposition of a real matrix in accordance with Theorem 9 enables us to obtain the fundamental formulas (107), (109), (111), and (113) without embedding the euclidean space in a unitary space, as was done above. This second approach to the fundamental formulas is based on the following theorem:

**THEOREM 10:** *If two real normal matrices are similar,*

$$B = T^{-1}AT \quad (AA^T = A^T A, BB^T = B^T B, A = \overline{A}, B = \overline{B}), \quad (127)$$

*then they are real-similar and orthogonally-similar:*

$$B = Q^{-1}AQ \quad (Q = \overline{Q} = Q^T)^{-1}. \quad (128)$$

*Proof:* Since the normal matrices  $A$  and  $B$  have the same characteristic values, there exists a polynomial  $g(\lambda)$  (see 2. on p. 272) such that

$$A^T = g(A), B^T = g(B).$$

Therefore the equation

$$g(B) = T^{-1}g(A)T,$$

which is a consequence of (127), can be written as follows:

$$B^T = T^{-1}A^T T. \quad (129)$$

When we go over to the transposed matrices in this equation, we obtain:

$$B = T^T A T^{T^{-1}}. \quad (130)$$

comparison of (127) with (130) shows that

$$T T^T A = A T T^T. \quad (131)$$

Now we make use of the polar decomposition of  $T$ :

$$T = SQ, \quad (132)$$

where  $S = \sqrt{TT^T} = h(TT^T)$  ( $h(\lambda)$  a polynomial) is symmetric and  $Q$  is real and orthogonal. Since  $A$ , by (131), is permutable with  $TT^T$ , it is also permutable with  $S = h(TT^T)$ . Therefore, when we substitute the expression for  $T$  from (132) in (127), we have:

$$B = Q^{-1}S^{-1}ASQ = Q^{-1}AQ.$$

This completes the proof.

Let us consider the real canonical matrix

$$\left\{ \left\| \begin{array}{cc} \mu_1 & \nu_1 \\ -\nu_1 & \mu_1 \end{array} \right\|, \dots, \left\| \begin{array}{cc} \mu_q & \nu_q \\ -\nu_q & \mu_q \end{array} \right\|, \mu_{2q+1}, \dots, \mu_n \right\}. \quad (133)$$

The matrix (133) is normal and has the characteristic values  $\mu_1 \pm i\nu_1, \dots, \mu_q \pm i\nu_q, \mu_{2q+1}, \dots, \mu_n$ . Since normal matrices are of simple structure, every normal matrix having the same characteristic values is similar (and by Theorem 10, real-similar and orthogonally-similar) to the matrix (133). Thus we arrive at the formula (107).

The formulas (109), (111), and (113) are obtained in exactly the same way.

### § 15. Commuting Normal Operators

In § 10 we have shown that two commuting operators  $A$  and  $B$  in an  $n$ -dimensional unitary space  $R$  always have a common characteristic vector. By mathematical induction we can show that this statement is true not only for two, but for any finite number, of commuting operators. For given  $m$  pairwise commuting operators  $A_1, A_2, \dots, A_m$  the first  $m-1$  of which have a common characteristic vector  $x$ , by repeating verbatim the argument of Lemma 1 (p. 270) (for  $A$  we take any  $A_i$  ( $i=1, 2, \dots, m-1$ ) and for  $B$  we take  $A_m$ ), we obtain a vector  $y$  which is a common characteristic vector of  $A_1, A_2, \dots, A_m$ .

This statement is even true for an infinite set of commuting operators, because such a set can only contain a finite number ( $\leq n^2$ ) of linearly independent operators, and a common characteristic value of the latter is a common characteristic value of all the operators of the given set.

**2.** Now suppose that an arbitrary finite or infinite set of pairwise commuting *normal* operators  $A, B, C, \dots$  is given. They all have a common characteristic vector  $x_1$ . We denote by  $T_1$  the  $(n-1)$ -dimensional sub-

space consisting of all vectors of  $\mathbf{R}$  that are orthogonal to  $\mathbf{x}_1$ . By § 10, 3. (p. 272), the subspace  $\mathbf{T}_1$  is invariant with respect to  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$ . Therefore all these operators have a common characteristic vector  $\mathbf{x}_2$  in  $\mathbf{T}_1$ . We consider the orthogonal complement  $\mathbf{T}_2$  of the plane  $[\mathbf{x}_1, \mathbf{x}_2]$  and select in it a vector  $\mathbf{x}_3$ , etc. Thus we obtain an orthogonal system  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  of common characteristic vectors of  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$ . These vectors can be normalized. Hence we have proved:

**THEOREM 11:** *If a finite or infinite set of pairwise commuting normal operators  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$  in a unitary space  $\mathbf{R}$  is given, then all these operators have a complete orthonormal system of common characteristic vectors  $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n$ :*

$$\mathbf{A}\mathbf{z}_i = \lambda_i \mathbf{z}_i, \quad \mathbf{B}\mathbf{z}_i = \lambda'_i \mathbf{z}_i, \quad \mathbf{C}\mathbf{z}_i = \lambda''_i \mathbf{z}_i, \quad \dots, \quad [(\mathbf{z}_i, \mathbf{z}_k) = \delta_{ik}; \quad i, k = 1, 2, \dots, n]. \quad (134)$$

In matrix form, this theorem reads as follows:

**THEOREM 11':** *If a finite or infinite set of pairwise commuting normal matrices  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$  is given, then all these matrices can be carried by one and the same unitary transformation into diagonal form, i.e., there exists a unitary matrix  $\mathbf{U}$  such that*

$$\left. \begin{aligned} \mathbf{A} &= \mathbf{U} \{ \lambda_1, \dots, \lambda_n \} \mathbf{U}^{-1}, \quad \mathbf{B} = \mathbf{U} \{ \lambda'_1, \dots, \lambda'_n \} \mathbf{U}^{-1}, \\ \mathbf{C} &= \mathbf{U} \{ \lambda''_1, \dots, \lambda''_n \} \mathbf{U}^{-1}, \dots \quad (\mathbf{U} = \mathbf{U}^{*-1}). \end{aligned} \right\} \quad (135)$$

Now suppose that commuting normal operators in a euclidean space  $\mathbf{R}$  are given. We denote by  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$  the linearly independent ones among them (their number is finite). We embed  $\mathbf{R}$  (under preservation of the metric) in a unitary space  $\tilde{\mathbf{R}}$ , as was done in § 13. Then by Theorem 11, the operators  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$  have a complete orthonormal system of common characteristic vectors  $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n$  in  $\tilde{\mathbf{R}}$ , i.e., (134) is satisfied.

We consider an arbitrary linear combination of  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$ :

$$\mathbf{P} = \alpha \mathbf{A} + \beta \mathbf{B} + \gamma \mathbf{C} + \dots$$

For arbitrary real values  $\alpha, \beta, \gamma, \dots$   $\mathbf{P}$  is a real ( $\mathbf{P}\mathbf{R} \subset \mathbf{R}$ ) normal operator in  $\tilde{\mathbf{R}}$  and

$$\left. \begin{aligned} \mathbf{P}\mathbf{z}_j &= \Lambda_j \mathbf{z}_j, \quad \Lambda_j = \alpha \lambda_j + \beta \lambda'_j + \gamma \lambda''_j + \dots \\ [(\mathbf{z}_j, \mathbf{z}_k) &= \delta_{jk}; \quad j, k = 1, 2, \dots, n]. \end{aligned} \right\} \quad (136)$$

The characteristic values  $\Lambda_j$  ( $j = 1, 2, \dots, n$ ) of  $\mathbf{P}$  are linear forms in  $\alpha, \beta, \gamma, \dots$ . Since  $\mathbf{P}$  is real, these forms can be split in pairs of complex conjugates and real ones; with a suitable numbering of the characteristic vectors, we have

$$A_{2k-1} = M_k + iN_k, \quad A_{2k} = M_k - iN_k, \quad A_l = M_l \quad (137)$$

$$(k = 1, 2, \dots, q; \quad l = 2q + 1, \dots, n),$$

where  $M_k$ ,  $N_k$ , and  $M_l$  are real linear forms in  $\alpha, \beta, \gamma, \dots$

We may assume that in (136) the corresponding vectors  $\mathbf{x}_{2k-1}$  and  $\mathbf{x}_{2k}$  are complex conjugates, and the  $\mathbf{x}_l$  real:

$$\mathbf{x}_{2k-1} = \mathbf{x}_k + i\mathbf{y}_k, \quad \mathbf{x}_{2k} = \mathbf{x}_k - i\mathbf{y}_k, \quad \mathbf{x}_l = \mathbf{x}_l \quad (138)$$

$$(k = 1, 2, \dots, q; \quad l = 2q + 1, \dots, n).$$

But then, as is easy to see, the real vectors

$$\mathbf{x}_k, \mathbf{y}_k, \mathbf{x}_l \quad (k = 1, 2, \dots, q; \quad l = 2q + 1, \dots, n) \quad (139)$$

form an orthonormal basis of  $\mathbf{R}$ . In this canonical basis we have:<sup>41</sup>

$$\begin{aligned} P\mathbf{x}_k &= M_k\mathbf{x}_k - N_k\mathbf{y}_k, \\ P\mathbf{y}_k &= N_k\mathbf{x}_k + M_k\mathbf{y}_k, \\ P\mathbf{x}_l &= M_l\mathbf{x}_l \end{aligned} \quad \left( \begin{array}{l} k = 1, 2, \dots, q, \\ l = 2q + 1, \dots, n \end{array} \right). \quad (140)$$

Since all the operators of the given set are obtained from  $\mathbf{P}$  for special values of  $\alpha, \beta, \gamma, \dots$  the basis (139), which does not depend on these parameters, is a common canonical basis for all the operators. Thus we have proved:

**THEOREM 12:** *If an arbitrary set of commuting normal linear operators in a euclidean space  $\mathbf{R}$  is given, then all these operators have a common orthonormal canonical basis  $\mathbf{x}_k, \mathbf{y}_k, \mathbf{x}_l$ :*

$$\left. \begin{aligned} A\mathbf{x}_k &= \mu_k\mathbf{x}_k - \nu_k\mathbf{y}_k, & B\mathbf{x}_k &= \mu'_k\mathbf{x}_k - \nu'_k\mathbf{y}_k, \dots, \\ A\mathbf{y}_k &= \nu_k\mathbf{x}_k + \mu_k\mathbf{y}_k, & B\mathbf{y}_k &= \nu'_k\mathbf{x}_k + \mu'_k\mathbf{y}_k, \dots, \\ A\mathbf{x}_l &= \mu_l\mathbf{x}_l; & B\mathbf{x}_l &= \mu'_l\mathbf{x}_l, \dots \end{aligned} \right\} \quad (141)$$

We give the matrix form of Theorem 12:

**THEOREM 12':** *Every set of commuting normal real matrices  $A, B, C, \dots$  can be carried by one and the same real orthogonal transformation  $Q$  into canonical form*

$$\left. \begin{aligned} A &= Q \left\{ \left\| \begin{array}{cc} \mu_1 & \nu_1 \\ -\nu_1 & \mu_1 \end{array} \right\|, \dots, \left\| \begin{array}{cc} \mu_q & \nu_q \\ -\nu_q & \mu_q \end{array} \right\|, \mu_{2q+1}, \dots, \mu_n \right\} Q^{-1}, \\ B &= Q \left\{ \left\| \begin{array}{cc} \mu'_1 & \nu'_1 \\ -\nu'_1 & \mu'_1 \end{array} \right\|, \dots, \left\| \begin{array}{cc} \mu'_q & \nu'_q \\ -\nu'_q & \mu'_q \end{array} \right\|, \mu'_{2q+1}, \dots, \mu'_n \right\} Q^{-1}, \\ &\dots \dots \dots \end{aligned} \right\} \quad (142)$$

<sup>41</sup> The equation (140) follows from (136), (137), and (138).

*Note.* If one of the operators  $A, B, C, \dots$  (matrices  $A, B, C, \dots$ )—say  $A$  ( $A$ )—is symmetric, then in the corresponding formulas (141) ((142)) all the  $\nu$  are zero. In the case of skew-symmetry, all the  $\mu$  are zero. In the case where  $A$  is an orthogonal operator ( $A$  an orthogonal matrix), we have  $\mu_k = \cos \varphi_k$ ,  $\nu_k = \sin \varphi_k$ ,  $\mu_l = \pm 1$  ( $k=1, 2, \dots, q$ ;  $l=2q+1, \dots, n$ ).



## CHAPTER X

### QUADRATIC AND HERMITIAN FORMS

#### § 1. Transformation of the Variables in a Quadratic Form

1. A *quadratic form* is a homogeneous polynomial of the second degree in  $n$  variables  $x_1, x_2, \dots, x_n$ . A quadratic form always has a representation

$$\sum_{i,k=1}^n a_{ik} x_i x_k \quad (a_{ik} = a_{ki}; i, k = 1, 2, \dots, n),$$

where  $A = \| a_{ik} \|_1^n$  is a symmetric matrix.

If we denote the column matrix  $(x_1, x_2, \dots, x_n)$  by  $x$  and denote the quadratic form by

$$A(x, x) = \sum_{i,k=1}^n a_{ik} x_i x_k, \quad (1)$$

then we can write:<sup>1</sup>

$$A(x, x) = x^T A x. \quad (2)$$

If  $A = \| a_{ik} \|_1^n$  is a real symmetric matrix, then the form (1) is called *real*. In this chapter we shall mainly be concerned with real quadratic forms.

The determinant  $|A| = |a_{ik}|_1^n$  is called the *discriminant* of the quadratic form  $A(x, x)$ . The form is called *singular* if its discriminant is zero.

To every quadratic form there corresponds a *bilinear form*

$$A(x, y) = \sum_{i,k=1}^n a_{ik} x_i y_k, \quad (3)$$

or

$$A(x, y) = x^T A y \quad (x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)). \quad (4)$$

If  $x^1, x^2, \dots, x^l, y^1, y^2, \dots, y^m$  are column matrices and  $c_1, c_2, \dots, c_l, d_1, d_2, \dots, d_m$  are scalars, then by the bilinearity of  $A(x, y)$  (see (4)),

<sup>1</sup> The sign  $\top$  denotes transposition. In (2) the quadratic form is represented as a product of three matrices: the row  $x^T$ , the square matrix  $A$ , and the column  $x$ .

$$A \left( \sum_{i=1}^l c_i x^i, \sum_{j=1}^m d_j y^j \right) = \sum_{i=1}^l \sum_{j=1}^m c_i d_j A(x^i, y^j). \quad (5)$$

If  $A$  is an operator in an  $n$ -dimensional euclidean space and if in some orthonormal basis  $e_1, e_2, \dots, e_n$  this symmetric operator corresponds to the matrix  $A = \|a_{ik}\|_1^n$ , then for arbitrary vectors

$$x = \sum_{i=1}^n x_i e_i, \quad y = \sum_{i=1}^n y_i e_i$$

we have the identity<sup>2</sup>

$$A(x, y) = (Ax, y) = (x, Ay).$$

In particular,

$$A(x, x) = (Ax, x) = (x, Ax),$$

where

$$a_{ik} = (Ae_i, e_k) \quad (i, k = 1, 2, \dots, n).$$

**2.** Let us see how the coefficient matrix of the form changes under a transformation of the variables:

$$x_i = \sum_{k=1}^n t_{ik} \xi_k \quad (i = 1, 2, \dots, n). \quad (6)$$

In matrix notation, this transformation looks as follows:

$$x = T\xi. \quad (6')$$

Here  $x, \xi$  are column matrices:  $x = (x_1, x_2, \dots, x_n)$  and  $\xi = (\xi_1, \xi_2, \dots, \xi_n)$ ; and  $T$  is the transforming matrix:  $T = \|t_{ik}\|_1^n$ .

Substituting the expression for  $x$  in (2), we obtain from (6'):

$$A(x, x) = \xi^T T^T A T \xi = \xi^T \tilde{A} \xi = \tilde{A}(\xi, \xi),$$

where

$$\tilde{A} = T^T A T. \quad (7)$$

The formula (7) expresses the coefficient matrix  $A = \|a_{ik}\|_1^n$  of the transformed form  $\tilde{A}(\xi, \xi) = \sum_{i,k=1}^n \tilde{a}_{ik} \xi_i \xi_k$  in terms of the coefficient matrix of the original form  $A = \|a_{ik}\|_1^n$  and the transformation matrix  $T = \|t_{ik}\|_1^n$ .

It follows from (7) that under a transformation the discriminant of the form is multiplied by the square of the determinant of the transformation:

<sup>2</sup> In  $A(x, y)$ , the parentheses form part of the notation; in  $(Ax, y)$  and  $(x, Ay)$ , they denote the scalar product.

$$|\tilde{A}| = |A| |T|^2. \quad (8)$$

In what follows we shall make use exclusively of non-singular transformations of the variables ( $|T| \neq 0$ ). Under such transformations, as is clear from (7), the rank of the coefficient matrix remains unchanged (the rank of  $A$  is the same as that of  $\tilde{A}$ ).<sup>3</sup> The rank of the coefficient matrix is usually called the *rank of the quadratic form*.

DEFINITION 1: Two symmetric matrices  $A$  and  $\tilde{A}$  connected as in formula (7), with  $|T| \neq 0$ , are called *congruent*.

Thus, a whole class of congruent symmetric matrices is associated with every quadratic form. As mentioned above, all these matrices have one and the same rank, the rank of the form. The rank is an invariant for the given class of matrices. In the real case, a second invariant is the so-called 'signature' of the quadratic form. We shall now proceed to introduce this concept.

## § 2. Reduction of a Quadratic Form to a Sum of Squares.

### The Law of Inertia

1. A real quadratic form  $A(x, x)$  can be represented in an infinite number of ways in the form

$$A(x, x) = \sum_{i=1}^r a_i X_i^2, \quad (9)$$

where  $a_i \neq 0$  ( $i = 1, 2, \dots, r$ ) and

$$X_i = \sum_{k=1}^n \alpha_{ki} x_k \quad (i = 1, 2, \dots, r)$$

are *linearly independent* real linear forms in the variables  $x_1, x_2, \dots, x_n$  (so that  $r \leq n$ ).

Let us consider a non-singular transformation of the variables under which the first  $r$  of the new variables  $\xi_1, \xi_2, \dots, \xi_n$  are connected with  $x_1, x_2, \dots, x_n$  by the formulas<sup>4</sup>

$$\xi_i = X_i \quad (i = 1, 2, \dots, r)$$

Then, in the new variables,

<sup>3</sup> See p. 17.

<sup>4</sup> We obtain the necessary transformation by adjoining to the system of linear forms  $X_1, \dots, X_r$  such linear forms  $X_{r+1}, \dots, X_n$  that the forms  $X_j$  ( $j = 1, 2, \dots, n$ ) are linearly independent and then setting  $\xi_j = X_j$  ( $j = 1, 2, \dots, n$ ).

$$A(x, x) = \tilde{A}(\xi, \xi) = \sum_{i=1}^r a_i \xi_i^2$$

and therefore  $\tilde{A} = \{a_1, a_2, \dots, a_r, 0, \dots, 0\}$ . But the rank of  $\tilde{A}$  is  $r$ . Hence: *The number of squares in the representation (9) is always equal to the rank of the form.*

2. We shall show that not only is the total number of squares invariant in the various representations of  $A(x, x)$  in the form (9), but also so is the number of positive<sup>5</sup> (and, hence, the number of negative) squares.

**THEOREM 1** (The Law of Inertia for Quadratic Forms): *In a representation of a real quadratic form  $A(x, x)$  as a sum of independent squares<sup>6</sup>*

$$A(x, x) = \sum_{i=1}^r a_i X_i^2, \quad (9)$$

*the number of positive and the number of negative squares are independent of the choice of the representation.*

*Proof.* Let us assume that we have, in addition to (9), another representation of  $A(x, x)$  in the form of a sum of independent squares

$$A(x, x) = \sum_{i=1}^r b_i Y_i^2$$

and that

$$\begin{aligned} a_1 > 0, a_2 > 0, \dots, a_g > 0, a_{g+1} < 0, \dots, a_r < 0, \\ b_1 > 0, b_2 > 0, \dots, b_h > 0, b_{h+1} < 0, \dots, b_r < 0, \end{aligned}$$

Suppose that  $g \neq h$ , say  $g < h$ . Then in the identity

$$\sum_{i=1}^r a_i X_i^2 = \sum_{i=1}^r b_i Y_i^2 \quad (10)$$

we give to the variables  $x_1, x_2, \dots, x_n$  values that satisfy the system of  $r - (h - g)$  equations

$$X_1 = 0, X_2 = 0, \dots, X_g = 0, Y_{h+1} = 0, \dots, Y_r = 0, \quad (11)$$

<sup>5</sup> By the number of positive (negative) squares in (9) we mean the number of positive (or negative)  $a_i$ .

<sup>6</sup> By a sum of independent squares we mean a sum of the form (9) in which all  $a_i \neq 0$  and the forms  $X_1, X_2, \dots, X_r$  are linearly independent.

and for which at least one of the forms  $X_{g+1}, \dots, X_r$  does not vanish.<sup>7</sup> For these values of the variables the left-hand side of the identity is

$$\sum_{j=g+1}^r a_j X_j^2 < 0,$$

and the right-hand side is

$$\sum_{k=1}^h b_k Y_k^2 \geq 0.$$

Thus, the assumption  $g \neq h$  has led to a contradiction, and the theorem is proved.

**DEFINITION 2:** *The difference  $\sigma$  between the number  $\pi$  of positive squares and the number  $\nu$  of negative squares in the representation of  $A(x, x)$  is called the signature of the form  $A(x, x)$ . (Notation:  $\sigma = \sigma[A(x, x)]$ ).*

The rank  $r$  and the signature  $\sigma$  determine the numbers  $\pi$  and  $\nu$  uniquely, since

$$r = \pi + \nu, \quad \sigma = \pi - \nu.$$

Note that in (9) the positive factor  $\sqrt{|a_i|}$  can be absorbed into the form  $X_i$  ( $i = 1, 2, \dots, r$ ). Then (9) assumes the form

$$A(x, x) = X_1^2 + X_2^2 + \dots + X_\pi^2 - X_{\pi+1}^2 - \dots - X_r^2. \quad (12)$$

Setting<sup>8</sup>  $\xi_i = X_i$  ( $i = 1, 2, \dots, r$ ), we reduce  $A(x, x)$  to the canonical form

$$\tilde{A}(\xi, \xi) = \xi_1^2 + \xi_2^2 + \dots + \xi_\pi^2 - \xi_{\pi+1}^2 - \dots - \xi_r^2. \quad (13)$$

Hence we deduce from Theorem 1 that: *Every real symmetric matrix  $A$  is congruent to a diagonal matrix in which the diagonal elements are  $+1$ ,  $-1$ , or  $0$ :*

$$A = T^T \left\{ \underbrace{+1, \dots, +1}_\pi, \underbrace{-1, \dots, -1}_\nu, 0, \dots, 0 \right\} T. \quad (14)$$

In the next section we shall give a rule for determining the signature from the coefficients of the quadratic form.

<sup>7</sup> Such values exist, since otherwise the equations  $X_{g+1} = 0, \dots, X_r = 0$  and hence all the equations  $X_1 = 0, X_2 = 0, \dots, X_r = 0$  would be consequences of the  $r - (h - g)$  equations (11). This is impossible, because the linear forms  $X_1, X_2, \dots, X_r$  are independent.

<sup>8</sup> See footnote 4.

### § 3. The Methods of Lagrange and Jacobi of Reducing a Quadratic Form to a Sum of Squares

It follows from the preceding section that in order to determine the rank and the signature of a form it is sufficient to reduce it in any way to a sum of independent squares.

We shall describe here two reduction methods: that of Lagrange and that of Jacobi.

**1. Lagrange's Method.** Let a quadratic form

$$A(x, x) = \sum_{i, k=1}^n a_{ik} x_i x_k$$

be given.

We consider two cases:

1) For some  $g$  ( $1 \leq g \leq n$ ) the diagonal coefficient  $a_{gg}$  is not equal to zero. Then we set

$$A(x, x) = \frac{1}{a_{gg}} \left( \sum_{k=1}^n a_{gk} x_k \right)^2 + A_1(x, x) \quad (15)$$

and convince ourselves by direct verification that the quadratic form  $A_1(x, x)$  does not contain the variable  $x_g$ . This method of separating out a square form in a quadratic form is always applicable when there is a non-zero diagonal element in the matrix  $A = \| a_{ik} \|_1^n$ .

2)  $a_{gg} = 0$  and  $a_{hh} = 0$ , but  $a_{gh} \neq 0$ . Then we set:

$$A(x, x) = \frac{1}{2a_{gh}} \left[ \sum_{k=1}^n (a_{gk} + a_{hk}) x_k \right]^2 - \frac{1}{2a_{gh}} \left[ \sum_{k=1}^n (a_{gk} - a_{hk}) x_k \right]^2 + A_2(x, x). \quad (16)$$

The forms

$$\sum_{k=1}^n a_{gk} x_k, \quad \sum_{k=1}^n a_{hk} x_k \quad (17)$$

are linearly independent, since the first contains  $x_h$  but not  $x_g$ , and the second contains  $x_g$  but not  $x_h$ . Therefore, in (16), the forms within the brackets are linearly independent (as sum and difference, respectively, of the independent linear forms (17)).

Therefore we have separated out two independent squares in  $A(x, x)$ . Each of these squares contains  $x_g$  and  $x_h$ , whereas  $A_2(x, x)$  does not contain these variables, as is easy to verify.

By successive application of a combination of the methods 1) and 2), we can always reduce the form  $A(x, x)$  by means of rational operations to a sum of squares. Moreover, the squares so obtained are linearly independent, since at each stage the square that is separated out contains an unknown that does not occur in the subsequent squares.

Note that the basic formulas (15) and (16) can be written as follows

$$A(x, x) = \frac{1}{4a_{gg}} \left( \frac{\partial A}{\partial x_g} \right)^2 + A_1(x, x), \quad (15')$$

$$A(x, x) = \frac{1}{8a_{gh}} \left[ \left( \frac{\partial A}{\partial x_g} + \frac{\partial A}{\partial x_h} \right)^2 - \left( \frac{\partial A}{\partial x_g} - \frac{\partial A}{\partial x_h} \right)^2 \right] + A_2(x, x). \quad (16')$$

*Example.*

$$A(x, x) = 4x_1^2 + x_2^2 + x_3^2 + x_4^2 - 4x_1x_2 - 4x_1x_3 + 4x_1x_4 + 4x_2x_3 - 4x_2x_4.$$

We apply formula (15') with  $g=1$ :

$$\begin{aligned} A(x, x) &= \frac{1}{16} (8x_1 - 4x_2 - 4x_3 + 4x_4)^2 + A_1(x, x) \\ &= (2x_1 - x_2 - x_3 + x_4)^2 + A_1(x, x), \end{aligned}$$

where

$$A_1(x, x) = 2x_2x_3 - 2x_2x_4 + 2x_3x_4.$$

We apply formula (16') with  $g=2$  and  $h=3$ :

$$\begin{aligned} A_1(x, x) &= \frac{1}{8} (2x_2 + 2x_3)^2 - \frac{1}{8} (2x_3 - 2x_2 - 4x_4)^2 + A_2(x, x) \\ &= \frac{1}{2} (x_2 + x_3)^2 - \frac{1}{2} (x_3 - x_2 - 2x_4)^2 + A_2(x, x), \end{aligned}$$

where

$$A_2(x, x) = 2x_4^2.$$

Finally,

$$\begin{aligned} A(x, x) &= (2x_1 - x_2 - x_3 + x_4)^2 + \frac{1}{2} (x_2 + x_3)^2 - \frac{1}{2} (x_3 - x_2 - 2x_4)^2 + 2x_4^2, \\ &= 4, \quad \sigma = 2. \end{aligned}$$

**2. Jacobi's Method.** We denote the rank of  $A(x, x) = \sum_{i,k=1}^n a_{ik}x_i x_k$  by  $r$  and assume that

$$D_k = A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \neq 0 \quad (k=1, 2, \dots, r).$$

Then the symmetric matrix  $A = \| a_{ik} \|_1^n$  can be reduced to the form

$$G = \begin{vmatrix} g_{11} & g_{12} & \dots & g_{1n} \\ 0 & g_{22} & \dots & g_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & g_{rr} & \dots & g_{rn} \\ 0 & 0 & \dots & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & 0 \end{vmatrix} \quad (18)$$

by Gauss's elimination algorithm (see Chapter II, § 1).

The elements of  $G$  are expressed in terms of the elements of  $A$  by the well-known formulas<sup>9</sup>

$$g_{pq} = \frac{A \begin{pmatrix} 1 & 2 & \dots & p-1 & p \\ 1 & 2 & \dots & p-1 & q \end{pmatrix}}{A \begin{pmatrix} 1 & 2 & \dots & p-1 \\ 1 & 2 & \dots & p-1 \end{pmatrix}} \quad (q = p, p+1, \dots, n; p = 1, 2, \dots, r). \quad (19)$$

In particular,

$$g_{pp} = \frac{D_p}{D_{p-1}} \quad (p = 1, 2, \dots, r; D_0 = 1). \quad (20)$$

In Chapter II, § 4 (formula (55) on page 41) we have shown that

$$A = G^T \hat{D} G, \quad (21)$$

where  $\hat{D}$  is the diagonal matrix:

$$\hat{D} = \left\{ \frac{1}{D_1}, \frac{D_1}{D_2}, \dots, \frac{D_{r-1}}{D_r}, 0, \dots, 0 \right\} = \left\{ \frac{1}{g_{11}}, \frac{1}{g_{22}}, \dots, \frac{1}{g_{rr}}, 0, \dots, 0 \right\}. \quad (22)$$

Without infringing (21) we may replace some of the zeros in the last  $n-r$  rows of  $G$  by arbitrary elements. By such a replacement we can make  $G$  into a non-singular upper triangular matrix

$$T = \begin{vmatrix} g_{11} & g_{12} & \dots & g_{1n} \\ 0 & g_{22} & \dots & g_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & g_{rr} & \dots & g_{rn} \\ 0 & 0 & \dots & 0 & * & \dots & * \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & \dots & * & \dots \end{vmatrix} \quad (|T| \neq 0). \quad (23)$$

<sup>9</sup> See Chapter II, § 2.



The equation (21) can then be rewritten:

$$A = T^T \hat{D} T. \quad (24)$$

From this equation it follows that the quadratic form<sup>10</sup>

$$\hat{D}(\xi, \xi) = \sum_{k=1}^r \frac{D_{k-1}}{D_k} \xi_k^2 = \sum_{k=1}^r \frac{\xi_k^2}{g_{kk}} \\ (\xi = (\xi_1, \xi_2, \dots, \xi_n); D_0 = 1)$$

goes over into the form  $A(x, x)$  under the transformation

$$\xi = Tx$$

Since

$$\xi_k = X_k, \quad X_k \equiv g_{kk}x_k + g_{k,k+1}x_{k+1} + \dots + g_{kn}x_n \quad (k=1, \dots, r), \quad (25)$$

we have *Jacobi's Formula*<sup>11</sup>

$$A(x, x) = \sum_{k=1}^r \frac{D_{k-1}}{D_k} X_k^2 = \sum_{k=1}^r \frac{X_k^2}{g_{kk}} \quad (D_0 = 1). \quad (26)$$

This formula gives a representation of  $A(x, x)$  in the form of a sum of  $r$  independent squares.<sup>12</sup>

Jacobi's formula is often given in another form.

Instead of  $X_k$  ( $k=1, 2, \dots, r$ ), the linearly independent forms

$$Y_k = D_{k-1}X_k \quad (k=1, 2, \dots, r; D_0=1) \quad (27)$$

are introduced. Then Jacobi's formula (26) can be written as:

$$A(x, x) = \sum_{k=1}^r \frac{Y_k^2}{D_{k-1}D_k}. \quad (28)$$

Here

$$Y_k = c_{kk}x_k + c_{k,k+1}x_{k+1} + \dots + c_{kn}x_n \quad (k=1, 2, \dots, r) \quad (29)$$

where

<sup>10</sup> We regard  $\hat{D}(\xi, \xi)$  as a quadratic form in the  $n$  variables  $\xi_1, \xi_2, \dots, \xi_n$ .

<sup>11</sup> Another approach to Jacobi's formula, which does not depend on (21), can be found, for example, in [17], pp. 43-44.

<sup>12</sup> The independence of the squares in Jacobi's formula follows from the fact that the form  $A(x, x)$  is of rank  $r$ . But we can also convince ourselves directly of the independence of the forms  $X_1, X_2, \dots, X_r$ . For, according to (20),  $g_{kk} = \frac{D_k}{D_{k-1}} \neq 0$  and therefore  $X_k$  contains the variable  $x_k$ , which does not occur in the forms  $X_{k+1}, \dots, X_r$  ( $k=1, 2, 3, \dots, r$ ). Hence  $X_1, X_2, \dots, X_r$  are linearly independent for all  $x$ .

$$c_{kq} = A \begin{pmatrix} 1 & 2 & \dots & k-1 & k \\ 1 & 2 & \dots & k-1 & q \end{pmatrix} \quad (q = k, k+1, \dots, n; k = 1, 2, \dots, r). \quad (30)$$

*Example.*

$$A(x, x) = x_1^2 + 3x_2^2 - 3x_3^2 - 4x_1x_2 + 2x_1x_3 - 2x_1x_4 - 6x_2x_3 + 8x_2x_4 + 2x_3x_4.$$

We reduce the matrix

$$A = \begin{vmatrix} 1 & -2 & 1 & -1 \\ -2 & 3 & -3 & 4 \\ 1 & -3 & 0 & 1 \\ -1 & 4 & 1 & -3 \end{vmatrix}$$

to the Gaussian form

$$G = \begin{vmatrix} 1 & -2 & 1 & -1 \\ 0 & -1 & -1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix}$$

Hence  $r = 2$ ,  $g_{11} = 1$ ,  $g_{22} = -1$ .

Jacobi's formula (26) yields:

$$A(x, x) = (x_1 - 2x_2 + x_3 - x_4)^2 - (-x_3 - x_3 + 2x_4)^2.$$

Jacobi's formula (28) yields the following theorem:

**THEOREM 2 (Jacobi).** *If for the quadratic form*

$$A(x, x) = \sum_{i,k=1}^n a_{ik} x_i x_k$$

*of rank  $r$  the inequality*

$$D_k = A \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \neq 0 \quad (k = 1, 2, \dots, r), \quad (31)$$

*holds, then the number  $\pi$  of positive squares and the number  $\nu$  of negative squares of  $A(x, x)$  coincide, respectively, with the number  $P$  of permanences of sign and the number  $V$  of variations of sign in the sequence*

$$1, D_1, D_2, \dots, D_r, \quad (32)$$

*i.e.,  $\pi = P(1, D_1, D_2, \dots, D_r)$ ,  $\nu = V(1, D_1, D_2, \dots, D_r)$ , and the signature*

$$\sigma = r - 2V(1, D_1, D_2, \dots, D_r). \quad (33)$$

Note 1. If in the sequence  $1, D_1, \dots, D_r \neq 0$  there are zeros, but not three in succession, then the signature can be determined by the use of the formula

$$\sigma = r - 2V(1, D_1, D_2, \dots, D_r)$$

omitting the zero  $D_k$  provided  $D_{k-1}D_{k+1} \neq 0$ , and setting

$$V(D_{k-1}, D_k, D_{k+1}, D_{k+2}) = \begin{cases} 1, & \text{when } \frac{D_{k+2}}{D_{k-1}} < 0, \\ 2, & \text{when } \frac{D_{k+2}}{D_{k-1}} > 0 \end{cases} \quad (34)$$

if  $D_k = D_{k+1} = 0$ .

We state this rule without proof.<sup>13</sup>

Note 2. When three consecutive zeros occur in  $D_1, D_2, \dots, D_{r-1}$ , then the signature of the quadratic form cannot be immediately determined by Jacobi's Theorem. In this case, the signs of the non-zero  $D_k$  do not determine the signature of the form. This is shown by the following example:

$$A(x, x) = 2a_1x_1x_4 + a_2x_2^2 + a_3x_3^2 \quad (a_1a_2a_3 \neq 0).$$

Here

$$D_1 = D_2 = D_3 = 0, \quad D_4 = -a_1^2a_2a_3 \neq 0.$$

But

$$\nu = \begin{cases} 1, & \text{when } a_2 > 0, a_3 > 0, \\ 3, & \text{when } a_2 < 0, a_3 < 0. \end{cases}$$

In both cases,  $D_4 < 0$ .

Note 3. If  $D_1 \neq 0, \dots, D_{r-1} \neq 0$ , but  $D_r = 0$ , then the signs of  $D_1, D_2, \dots, D_{r-1}$  do not determine the signature of the form. As a corroborating example, we can take the form

$$ax_1^2 + ax_2^2 + bx_3^2 + 2ax_1x_2 + 2ax_2x_3 + 2ax_1x_3 = a(x_1 + x_2 + x_3)^2 + (b-a)x_3^2.$$

#### § 4. Positive Quadratic Forms

1. In this section we deal with the special, but important, class of positive quadratic forms.

DEFINITION 3: A real quadratic form  $A(x, x) = \sum_{i,k=1}^n a_{ik}x_ix_k$  is called positive (negative) semidefinite if for arbitrary real values of the variables:

$$A(x, x) \geq 0 \quad (\leq 0). \quad (35)$$

<sup>13</sup> This rule was found in the case of a single zero  $D_k$  by Gundenfinger and for two successive zeros  $D_k$  by Frobenius [162].

DEFINITION 4: A real quadratic form  $A(x, x) = \sum_{i,k=1}^n a_{ik}x_i x_k$  is called *positive (negative) definite* if for arbitrary values of the variables, not all zero, ( $x \neq 0$ )

$$A(x, x) > 0 \quad (< 0). \quad (36)$$

The class of positive (negative) definite forms is part of the class of positive (negative) semidefinite forms.

Let  $A(x, x)$  be a positive-semidefinite form. We represent it in the form of a sum of linearly independent squares:

$$A(x, x) = \sum_{i=1}^r a_i X_i^2, \quad (37)$$

In this representation, all the squares must be positive:

$$a_i > 0 \quad (i = 1, 2, \dots, r). \quad (38)$$

For if any  $a_i$  were negative, then we could select values of  $x_1, x_2, \dots, x_n$  for which

$$X_1 = \dots = X_{i-1} = X_{i+1} = \dots = X_r = 0, \quad X_i \neq 0.$$

But then  $A(x, x)$  would have a negative value for these values of the variables, and by assumption this is impossible. It is clear that, conversely, it follows from (37) and (38) that the form  $A(x, x)$  is positive semidefinite.

Thus, a positive semidefinite quadratic form is characterized by the equations  $\sigma = r$  ( $\pi = r, \nu = 0$ ).

Now let  $A(x, x)$  be a positive-definite form. Then  $A(x, x)$  is also positive semidefinite. Therefore it is representable in the form (37), where all the  $a_i$  ( $i = 1, 2, \dots, r$ ) are positive. From the positive definiteness it follows that  $r = n$ . For if  $r < n$ , we could find values of  $x_1, x_2, \dots, x_n$ , not all zero, such that all the  $X_i$  would be zero. But then by (37)  $A(x, x) = 0$  for  $x \neq 0$ , and this contradicts (36).

It is easy to see that, conversely, if in (37)  $r = n$  and all the  $a_1, a_2, \dots, a_n$  are positive, then  $A(x, x)$  is a positive-definite form.

In other words: *A positive-semidefinite form is positive definite if and only if it is not singular.*

**2.** The following theorem gives a criterion for positive definiteness in the form of inequalities which the coefficients of the form must satisfy. We shall use the notation of the preceding section for the sequence of the principal minors of  $A$ :

$$D_1 = a_{11}, \quad D_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \quad \dots, \quad D_n = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

THEOREM 3: A quadratic form is positive definite if and only if

$$D_1 > 0, D_2 > 0, \dots, D_n > 0. \quad (39)$$

*Proof.* The sufficiency of the conditions (39) follows immediately from Jacobi's formula (28). The necessity of (39) is established as follows. From the fact that  $A(x, x) = \sum_{i,k=1}^n a_{ik}x_i x_k$  is positive definite, it follows that the 'restricted' forms<sup>14</sup>

$$A_p(x, x) = \sum_{i,k=1}^p a_{ik}x_i x_k \quad (p=1, 2, \dots, n)$$

are also positive definite. But then all these forms must be non singular, i.e.,

$$D_p = |A_p| \neq 0 \quad (p=1, 2, \dots, n).$$

We are now in a position to apply Jacobi's formula (28) (for  $r=n$ ). Since all the squares on the right-hand side of the formula must be positive, we have

$$D_1 > 0, D_1 D_2 > 0, D_1 D_2 D_3 > 0, \dots, D_{n-1} D_n > 0.$$

Hence the inequality (39) follows, and the theorem is proved.

Since every principal minor of  $A$  can be brought into the top left corner by a suitable numbering of the variables, we have the

COROLLARY: In a positive-definite quadratic form  $A(x, x) = \sum_{i,k=1}^n a_{ik}x_i x_k$ , all the principal minors of the coefficient matrix are positive:<sup>15</sup>

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} > 0 \quad (1 \leq i_1 < i_2 < \dots < i_p \leq n; p=1, 2, \dots, n).$$

*Note.* If the successive principal minors are non-negative,

$$D_1 \geq 0, D_2 \geq 0, \dots, D_n \geq 0, \quad (40)$$

<sup>14</sup> The form  $A_p(x, x)$  is obtained from  $A(x, x)$  if we set in the latter

$$x_{p+1} = \dots = x_n = 0 \quad (p=1, 2, \dots, n).$$

<sup>15</sup> Thus, when the successive principal minors of a real symmetric matrix are positive, all the remaining principal minors are then also positive.

it does *not* follow that  $A(x, x)$  is positive semidefinite. For, the form

$$a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

in which  $a_{11} = a_{12} = 0$ ,  $a_{22} < 0$  satisfies (40), but is not positive semidefinite.

However, we have the following theorem.

**THEOREM 4:** *A quadratic form  $A(x, x) = \sum_{i,k=1}^n a_{ik}x_ix_k$  is positive semidefinite if and only if all the principal minors of its coefficient matrix are non-negative:*

$$A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} \geq 0 \quad (1 \leq i_1 < i_2 < \dots < i_p \leq n; p = 1, 2, \dots, n). \quad (41)$$

*Proof.* We introduce the auxiliary form

$$A_\varepsilon(x, x) = A(x, x) + \varepsilon \sum_{i=1}^n x_i^2 \quad (\varepsilon < 0).$$

Obviously  $\lim_{\varepsilon \rightarrow 0} A_\varepsilon(x, x) = A(x, x)$ .

The fact that  $A(x, x)$  is positive semidefinite implies that  $A_\varepsilon(x, x)$  is positive definite, so that we have the inequality (cf. Corollary to Theorem 3)

$$A_\varepsilon \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} > 0 \quad (1 \leq i_1 < i_2 < \dots < i_p \leq n; p = 1, 2, \dots, n).$$

Proceeding to the limit for  $\varepsilon \rightarrow 0$ , we obtain (41).

Suppose, conversely, that (41) holds. Then we have

$$A_\varepsilon \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} = \varepsilon^p + \dots \geq \varepsilon^p > 0 \quad (1 \leq i_1 < i_2 < \dots < i_p \leq n; p = 1, 2, \dots, n).$$

But then (by Theorem 3),  $A_\varepsilon(x, x)$  is positive definite

$$A_\varepsilon(x, x) > 0 \quad (x \neq 0).$$

Proceeding to the limit for  $\varepsilon \rightarrow 0$  we obtain:

$$A(x, x) \geq 0.$$

This completes the proof.

The conditions for a form to be negative semidefinite and negative definite are obtained from (39) and (41), respectively, when these inequalities are applied to  $-A(x, x)$ .

**THEOREM 5:** *A quadratic form  $A(x, x)$  is negative definite if and only if the following inequalities hold:*

$$D_1 < 0, D_2 > 0, D_3 < 0, \dots, (-1)^n D_n > 0. \quad (42')$$

**THEOREM 6:** *A quadratic form  $A(x, x)$  is negative semidefinite if and only if the following inequalities hold:*

$$(-1)^p A \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} \geq 0 \quad (1 \leq i_1 < i_2 < \dots < i_p \leq n; p = 1, 2, \dots, n). \quad (43)$$

## § 5. Reduction of a Quadratic Form to Principal Axes

1. We consider an arbitrary real quadratic form

$$A(x, x) = \sum_{i, k=1}^n a_{ik} x_i x_k.$$

Its coefficient matrix  $A = \|a_{ik}\|_1^n$  is real and symmetric. Therefore (see Chapter IX, § 13) it is orthogonally similar to a real diagonal matrix  $\Lambda$ , i.e., there exists a real orthogonal matrix  $Q$  such that

$$\Lambda = Q^{-1} A Q \quad (\Lambda = \|\lambda_i \delta_{ik}\|_1^n, \quad Q Q^T = E). \quad (44)$$

Here  $\lambda_1, \lambda_2, \dots, \lambda_n$  are the characteristic values of  $A$ .

Since for an orthogonal matrix  $Q^{-1} = Q^T$ , it follows from (43) that under the orthogonal transformation of the variables

$$x = Q\xi \quad (Q Q^T = E) \quad (45)$$

or, in greater detail,

$$x_i = \sum_{k=1}^n q_{ik} \xi_k \quad \left( \sum_{i=1}^n q_{ij} q_{ik} = \delta_{jk}; j, k = 1, 2, \dots, n \right), \quad (45')$$

the form  $A(x, x)$  goes over into

$$A(\xi, \xi) = \sum_{i=1}^n \lambda_i \xi_i^2. \quad (46)$$

**THEOREM 7:** Every real quadratic form  $A(x, x) = \sum_{i,k=1}^n a_{ik}x_i x_k$  can be reduced to the canonical form (46) by an orthogonal transformation, where  $\lambda_1, \lambda_2, \dots, \lambda_n$  are the characteristic values of  $A = \|a_{ik}\|$ .

The reduction of the quadratic form  $A(x, x)$  to the canonical form (46) is called *reduction to principal axes*. The reason for this name is that the equation of a central hypersurface of the second order

$$\sum_{i,k=1}^n a_{ik}x_i x_k = c \quad (c = \text{const.} \neq 0) \quad (47)$$

under the orthogonal transformation (45') of the variables assumes the canonical form

$$\sum_{i=1}^n \varepsilon_i \frac{\xi_i^2}{a_i^2} = 1 \quad \left( \frac{\varepsilon_i}{a_i^2} = \frac{\lambda_i}{c}; \quad \varepsilon_i = \pm 1; \quad i = 1, 2, \dots, n \right). \quad (48)$$

If we regard  $x_1, x_2, \dots, x_n$  as coordinates in an orthonormal basis in an  $n$ -dimensional euclidean space, then  $\xi_1, \xi_2, \dots, \xi_n$  are the coordinates in a new orthonormal basis of the same space, and the 'rotation'<sup>16</sup> of the axes is brought about by the orthogonal transformation (45): The new coordinate axes are axes of symmetry of the central surface (47) and are usually called its *principal axes*.

**2.** It follows from (46) that the rank  $r$  of  $A(x, x)$  is equal to the number of non-zero characteristic values of  $A$  and the signature  $\sigma$  is equal to the difference between the number of positive and the number of negative characteristic values of  $A$ .

Hence, in particular, we have the following proposition:

*If under a continuous change of the coefficients of a quadratic form the rank remains unchanged, then the signature also remains unchanged.*

Here we have started from the fact that a continuous change of the coefficients produces a continuous change of the characteristic values. The signature can only change when some characteristic value changes sign. But then at some intermediate stage this characteristic value must pass through zero, and this results in a change of the rank of the form.

<sup>16</sup> If  $|Q| = -1$ , then (45) is a combination of a rotation with a reflection (see p. 287). However, the reduction to principal axes can always be effected by a proper orthogonal matrix ( $|Q| = 1$ ). This follows from the fact that, without changing the canonical form, we can perform the additional transformation

$$\xi_i = \xi_i^T \quad (i = 1, 2, \dots, n-1), \quad \xi_n = -\xi_n^T.$$



## § 6. Pencils of Quadratic Forms

1. In the theory of small oscillations it is necessary to consider simultaneously two quadratic forms one of which gives the potential, and the other the kinetic energy, of the system. The second form is always positive definite.

The study of a system of two such forms is the object of this section.

Two real quadratic forms

$$A(x, x) = \sum_{i, k=1}^n a_{ik} x_i x_k \quad \text{and} \quad B(x, x) = \sum_{i, k=1}^n b_{ik} x_i x_k$$

determine the *pencil* of forms  $A(x, x) - \lambda B(x, x)$  ( $\lambda$  is a parameter).

If the form  $B(x, x)$  is positive definite, the pencil  $A(x, x) - \lambda B(x, x)$  is then called *regular*.

The equation

$$|A - \lambda B| = 0$$

is called the *characteristic equation of the pencil of forms*  $A(x, x) - \lambda B(x, x)$ .

We denote by  $\lambda_0$  some root of this equation. Since the matrix  $A - \lambda_0 B$  is singular, there exists a column  $z = (z_1, z_2, \dots, z_n) \neq 0$  such that  $(A - \lambda_0 B)z = 0$ , or

$$Az = \lambda_0 Bz \quad (z \neq 0).$$

The number  $\lambda_0$  will be called a *characteristic value of the pencil*  $A(x, x) - \lambda B(x, x)$  and  $z$  a corresponding *principal column* or '*principal vector*' of the pencil. The following theorem holds:

**THEOREM 8:** *The characteristic equation*

$$|A - \lambda B| = 0$$

*of a regular pencil of forms*  $A(x, x) - \lambda B(x, x)$  *always has*  $n$  *real roots*  $\lambda_k$  *with the corresponding principal vectors*  $z^k = (z_{1k}, z_{2k}, \dots, z_{nk})$  ( $k = 1, 2, \dots, n$ ):

$$Az^k = \lambda_k Bz^k \quad (k = 1, 2, \dots, n). \quad (49)$$

*These principal vectors*  $z^k$  *can be chosen such that the relations*

$$B(z^i, z^k) = \delta_{ik} \quad (i, k = 1, 2, \dots, n) \quad (50)$$

*are satisfied.*

*Proof.* We observe that (49) can be written as:

$$B^{-1}Az^k = \lambda_k z^k \quad (k=1, 2, \dots, n). \quad (51)$$

Thus, our theorem states that the matrix

$$D = B^{-1}A \quad (52)$$

1. has simple structure, 2. has real characteristic values, and 3. has characteristic columns (vectors)  $z^1, z^2, \dots, z^n$  corresponding to these characteristic values and satisfying the relations (50).<sup>17</sup>

In order to prove these three statements, we introduce an  $n$ -dimensional vector space  $\mathbf{R}$  over the field of real numbers. In this space we fix a basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  and introduce a scalar product of two arbitrary vectors

$$\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i, \quad \mathbf{y} = \sum_{i=1}^n y_i \mathbf{e}_i$$

by means of the positive-definite bilinear form  $B(\mathbf{x}, \mathbf{y})$ :

$$(\mathbf{x}\mathbf{y}) = B(\mathbf{x}, \mathbf{y}) = \sum_{i,k=1}^n b_{ik} x_i y_k = \mathbf{x}^T B \mathbf{y} \quad (53)$$

and hence the square of the length of a vector  $\mathbf{x}$  by means of the form  $B(\mathbf{x}, \mathbf{x})$ :

$$(\mathbf{x}\mathbf{x}) = B(\mathbf{x}, \mathbf{x}) = \mathbf{x}^T B \mathbf{x}, \quad (53')$$

where  $\mathbf{x}$  and  $\mathbf{y}$  are columns  $\mathbf{x} = (x_1, x_2, \dots, x_n)$ ,  $\mathbf{y} = (y_1, y_2, \dots, y_n)$ .

It is easy to verify that the metric so introduced satisfies the postulates 1.-5. (p. 243) and is, therefore, euclidean.

We have obtained an  $n$ -dimensional euclidean space  $\mathbf{R}$ , but the original basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  is, in general, not orthonormal. To the matrices  $A, B$ , and  $D = B^{-1}A$  there correspond in this basis linear operators in  $\mathbf{R}$ :  $\mathbf{A}$ ,  $\mathbf{B}$ , and  $\mathbf{D} = \mathbf{B}^{-1}\mathbf{A}$ .<sup>18</sup>

<sup>17</sup> If  $D$  were a symmetric matrix, then the properties 1. and 2. would follow immediately from properties of a symmetric operator (Chapter IX, p. 284). However,  $D$ , as a product of two symmetric matrices, is not necessarily itself symmetric, since  $D = B^{-1}A$  and  $D^T = AB^{-1}$ .

<sup>18</sup> Since the basis  $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$  is not orthonormal, the operators  $\mathbf{A}$  and  $\mathbf{B}$  to which, in this basis, the symmetric matrices  $A$  and  $B$  correspond, are not necessarily symmetric themselves.

We shall show that  $D$  is a symmetric operator in  $R$  (see Chapter IX, § 13).<sup>19</sup> Indeed, for arbitrary vectors  $x$  and  $y$  with the coordinate columns  $x = (x_1, x_2, \dots, x_n)$  and  $y = (y_1, y_2, \dots, y_n)$  we have, by (52) and (53),

$$(Dx, y) = (Dx)^T By = x^T D^T By = x^T A B^{-1} By = x^T Ay$$

and

$$(x, Dy) = x^T B Dy = x^T B B^{-1} Ay = x^T Ay,$$

i.e.,

$$(Dx, y) = (x, Dy).$$

The symmetric operator  $D = B^{-1}A$  has real characteristic values  $\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$  and a complete orthonormal system of characteristic vectors  $z^1, z^2, z^3, \dots, z^n$  (see p. 284, Chapter IX) :

$$B^{-1}A z^k = \lambda_k z^k \quad (k = 1, 2, \dots, n), \quad (54)$$

$$(z^i, z^k) = \delta_{ik} \quad (i, k = 1, 2, \dots, n). \quad (54')$$

Let  $z^k = (z_{1k}, z_{2k}, \dots, z_{nk})$  be the coordinate column of  $z^k$  ( $k = 1, 2, \dots, n$ ) in the basis  $e_1, e_2, \dots, e_n$ . Then the equations (54) can be written in the form (51) or (49) and the relations (54'), by (53), yield the equation (50).

This completes the proof.

Note that it follows from (50) that the columns  $z^1, z^2, \dots, z^n$  are linearly independent. For suppose that

$$\sum_{k=1}^n c_k z^k = 0. \quad (55)$$

Then for every  $i$  ( $1 \leq i \leq n$ ), by (50),

$$0 = B \left( z^i, \sum_{k=1}^n c_k z^k \right) = \sum_{k=1}^n c_k B(z^i, z^k) = c_i.$$

Then all the  $c_i$  ( $i = 1, 2, \dots, n$ ) in (55) are zero and there is no linear dependence among the columns  $z^1, z^2, \dots, z^n$ .

A square matrix formed from principal columns  $z^1, z^2, \dots, z^n$  satisfying the relations (50)

$$Z = (z^1, z^2, \dots, z^n) = \|z_{ik}\|_1^n$$

will be called a *principal matrix* for the pencil of forms  $A(x, x) - \lambda B(x, x)$ .

<sup>19</sup> Hence  $D$  is similar to some symmetric matrix.

The principal matrix  $Z$  is non-singular ( $|Z| \neq 0$ ), because its columns are linearly independent.

The equation (50) can be written as follows:

$$z^i{}^\top Bz^k = \delta_{ik} \quad (i, k = 1, 2, \dots, n). \quad (56)$$

Moreover, when we multiply both sides of (49) on the left by the row matrix  $z^i{}^\top$ , we obtain:

$$z^i{}^\top A z^k = \lambda_k z^i{}^\top B z^k = \lambda_k \delta_{ik} \quad (i, k = 1, 2, \dots, n). \quad (57)$$

By introducing the principal matrix  $Z = (z^1, z^2, \dots, z^n)$ , we can represent (56) and (57) in the form

$$Z^\top A Z = \|\lambda_k \delta_{ik}\|_1^n, \quad Z^\top B Z = E. \quad (58)$$

The formulas (58) show that the non-singular transformation

$$x = Z\xi \quad (59)$$

reduces the quadratic forms  $A(x, x)$  and  $B(x, x)$  simultaneously to sums of squares:

$$\sum_{k=1}^n \lambda_k \xi_k^2 \quad \text{and} \quad \sum_{k=1}^n \xi_k^2. \quad (60)$$

This property of (59) characterizes a principal matrix  $Z$ . For suppose that the transformation (59) reduces the forms  $A(x, x)$  and  $B(x, x)$  simultaneously to the canonical forms (60). Then (58) holds, and hence (56) and (57) holds for  $Z$ . (58) implies that  $Z$  is non-singular ( $|Z| \neq 0$ ). We rewrite (57) as follows:

$$z^i{}^\top (A z^k - \lambda_k B z^k) = 0 \quad (i = 1, 2, \dots, n), \quad (61)$$

where  $k$  has an arbitrary fixed value ( $1 \leq k \leq n$ ). The system of equations (61) can be contracted into the single equation

$$Z^\top (A z_k - \lambda_k B z_k) = 0;$$

hence, since  $Z^\top$  is non-singular,

$$A z^k - \lambda_k B z^k = 0;$$

i.e., for every  $k$  (49) holds. Therefore  $Z$  is a principal matrix. Thus we have proved the following theorem:

THEOREM 9: If  $Z = \| z_{ik} \|_1^n$  is a principal matrix of a regular pencil of forms  $A(x, x) - \lambda B(x, x)$ , then the transformation

$$x = Z\xi \quad (62)$$

reduces the forms  $A(x, x)$  and  $B(x, x)$  simultaneously to sums of squares

$$\sum_{k=1}^n \lambda_k \xi_k^2, \quad \sum_{k=1}^n \xi_k^2, \quad (63)$$

where  $\lambda_1, \lambda_2, \dots, \lambda_n$  are the characteristic values of the pencil  $A(x, x) - \lambda B(x, x)$  corresponding to the columns  $z^1, z^2, \dots, z^n$  of  $Z$ .

Conversely, if some transformation (62) simultaneously reduces  $A(x, x)$  and  $B(x, x)$  to the form (63), then  $Z = \| z_{ik} \|_1^n$  is a principal matrix of the regular pencil of forms  $A(x, x) - \lambda B(x, x)$ .

Sometimes the characteristic property of the transformation (62) formulated in Theorem 9 is used for the construction of a principal matrix and the proof of Theorem 8.<sup>20</sup> For this purpose, we first of all carry out a transformation of variables  $x = Ty$  that reduces the form  $B(x, x)$  to the 'unit' sum of squares  $\sum_{k=1}^n y_k^2$  (which is always possible, since  $B(x, x)$  is positive definite). Then  $A(x, x)$  is carried into a certain form  $A_1(y, y)$ . Now the form  $A_1(y, y)$  is reduced to the form  $\sum_{k=1}^n \lambda_k \xi_k^2$  by an orthogonal transformation  $y = Q\xi$  (reduction to principal axes!). Then, obviously,<sup>21</sup>  $\sum_{k=1}^n y_k^2 = \sum_{k=1}^n \xi_k^2$ . Thus the transformation  $x = Z\xi$ , where  $Z = TQ$ , reduces the two given forms to (63). Afterwards it turns out (as we have shown on p. 313) that the columns  $z^1, z^2, \dots, z^n$  of  $Z$  satisfy the relations (49) and (50).

In the special case where  $B(x, x)$  is the unit form, i.e.,  $B(x, x) = \sum_{k=1}^n x_k^2$ , so that  $B = E$ , the characteristic equation of the pencil  $A(x, x) - \lambda B(x, x)$  coincides with the characteristic equation of  $A$ , and the principal vectors of the pencil are characteristic vectors of  $A$ . In this case the relations (50) can be written as follows:

$$z^i \tau z^k = \delta_{ik} \quad (i, k = 1, 2, \dots, n)$$

and they express the orthonormality of the columns  $z^1, z^2, \dots, z^n$ .

<sup>20</sup> See [17], pp. 56-57.

<sup>21</sup> An orthogonal transformation does not alter a sum of squares of the variables, because  $(Qx)^\tau Qx = x^\tau x$ .

2. Theorems 8 and 9 admit of an intuitive geometric interpretation. We introduce a euclidean space  $\mathbf{R}$  with the basis  $e_1, e_2, \dots, e_n$  and the fundamental metric form  $B(x, x)$  just as was done for the proof of Theorem 8. In  $\mathbf{R}$  we consider a central hypersurface of the second order whose equation is .

$$A(x, x) = \sum_{i,k=1}^n a_{ik} x_i x_k = c. \quad (64)$$

After the coordinate transformation  $x = Z\xi$ , where  $Z = \|z_{ik}\|_1^n$  is a principal matrix of the pencil  $A(x, x) - \lambda B(x, x)$ , the new basis vectors are the vectors  $z^1, z^2, \dots, z^n$  whose coordinates in the old basis form the columns of  $Z$ , i.e., the principal vectors of the pencil. These vectors form an orthonormal basis in which the equation of the hypersurface (64) has the form

$$\sum_{k=1}^n \lambda_k \xi_k^2 = c. \quad (65)$$

Therefore the principal vectors  $z^1, z^2, \dots, z^n$  of the pencil coincide in direction with the principal axes of the hypersurface (64), and the characteristic values  $\lambda_1, \lambda_2, \dots, \lambda_n$  of the pencil determine the lengths of the semi-axes:

$$\lambda_k = \pm \frac{c}{a_k^2} \quad (k=1, 2, \dots, n).$$

Thus, the task of determining the characteristic values and the principal vectors of a regular pencil of forms  $A(x, x) - \lambda B(x, x)$  is equivalent to the task of reducing the equation (64) of a central hypersurface of the second order to principal axes, provided the equation of the hypersurface is given in a general skew coordinate system<sup>22</sup> in which the 'unit sphere' has the equation  $B(x, x) = 1$ .

*Example.* Given the equation of a surface of the second order

$$2x^2 - 2y^2 - 3z^2 - 10yz + 2xz - 4 = 0 \quad (66)$$

in a general skew coordinate system in which the equation of the unit sphere is

$$2x^2 + 3y^2 + 2z^2 + 2xz = 1, \quad (67)$$

it is required to reduce equation (66) to principal axes.

In this case

$$A = \left\| \begin{array}{ccc} 2 & 0 & 1 \\ 0 & -2 & -5 \\ 1 & -5 & -3 \end{array} \right\|, \quad B = \left\| \begin{array}{ccc} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{array} \right\|.$$

<sup>22</sup> I.e., a skew coordinate system with distinct units of lengths along the axes.

The characteristic equation of the pencil  $|A - \lambda B| = 0$  has the form

$$\begin{vmatrix} 2-2\lambda & 0 & 1-\lambda \\ 0 & -2-3\lambda & -5 \\ 1-\lambda & -5 & -3-2\lambda \end{vmatrix} = 0. \quad (68)$$

This equation has three roots:  $\lambda_1 = 1$ ,  $\lambda_2 = 1$ ,  $\lambda_3 = -4$ .

We denote the coordinates of a principal vector corresponding to the characteristic value 1 by  $u, v, w$ . The values of  $u, v, w$  are determined from the system of homogeneous equations whose coefficients are the elements of the determinant (68) for  $\lambda = 1$ :

$$0 \cdot u + 0 \cdot v + 0 \cdot w = 0,$$

$$0 \cdot u - 5v - 5w = 0,$$

$$0 \cdot u - 5v - 5w = 0.$$

In fact we have only one relation

$$v + w = 0.$$

To the characteristic value  $\lambda = 1$  there must correspond two orthonormal principal vectors. The coordinates of the first can be chosen arbitrarily, provided they satisfy the relation  $v + w = 0$ .

We set

$$u = 0, v, w = -v.$$

We take the coordinates of the second principal vector in the form

$$u', v', w' = -v'$$

and write down the condition for orthogonality ( $B(z^1, z^2) = 0$ ):

$$2uu' + 3vv' + 2ww' + uw' + u'w = 0.$$

Hence we find:  $u' = 5v'$ . Thus, the coordinates of the second principal vector are

$$u' = 5v', v', w' = -v'.$$

Similarly, by setting  $\lambda = -4$  in the characteristic determinant, we find for the corresponding principal vector:

$$u'', v'' = -u'', w'' = -2u''.$$

The values of  $v$ ,  $v'$ , and  $u''$  are determined from the condition that the coordinates of a principal vector must satisfy the equation of the unit sphere ( $B(x, x) = 1$ ), i.e., (67). Hence we find:

$$v = \frac{1}{\sqrt{5}}, \quad v' = \frac{1}{3\sqrt{5}}, \quad u'' = -\frac{1}{3}.$$

Therefore the principal matrix has the form

$$Z = \begin{vmatrix} 0 & \frac{\sqrt{5}}{3} & -\frac{1}{3} \\ \frac{1}{\sqrt{5}} & \frac{1}{3\sqrt{5}} & \frac{1}{3} \\ -\frac{1}{\sqrt{5}} & -\frac{1}{3\sqrt{5}} & \frac{2}{3} \end{vmatrix},$$

and the corresponding coordinate transformation ( $x = Z\xi$ ) reduces the equations (66) and (67) to the canonical form

$$\xi_1^2 + \xi_2^2 - 4\xi_3^2 - 4 = 0, \quad \xi_1^2 + \xi_2^2 + \xi_3^2 = 1$$

The first equation can also be written as follows:

$$\frac{\xi_1^2}{4} + \frac{\xi_2^2}{4} - \frac{\xi_3^2}{1} = 1.$$

This is the equation of a one-sheet hyperboloid of rotation with real semi-axes equal to 2, and an imaginary one equal to 1. The coordinates of the endpoint of the axis of rotation is determined by the third column of  $Z$ , i.e.,  $-1/3, 1/3, 2/3$ . The coordinates of the endpoints of the other two orthogonal axes are given by the first and second columns.

## § 7. Extremal Properties of the Characteristic Values of a Regular Pencil of Forms<sup>23</sup>

I. Suppose that two quadratic forms are given

$$A(x, x) = \sum_{i, k=1}^n a_{ik} x_i x_k \quad \text{and} \quad B(x, x) = \sum_{i, k=1}^n b_{ik} x_i x_k,$$

of which  $B(x, x)$  is positive definite. We number the characteristic values of the regular pencil of forms  $A(x, x) - \lambda B(x, x)$  in non-descending order:

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n. \quad (69)$$

<sup>23</sup> In the exposition of this section, we follow the book [17], § 10.



The principal vectors<sup>24</sup> corresponding to these characteristic values are denoted, as before, by  $z^1, z^2, \dots, z^n$ :

$$z^k = (z_{1k}, z_{2k}, \dots, z_{nk}) \quad (k = 1, 2, \dots, n).$$

Let us determine the least value (minimum) of the ratio of the forms  $\frac{A(x, x)}{B(x, x)}$  considering all possible values of the variables, not all equal to zero ( $x \neq 0$ ). For this purpose it is convenient to go over to new variables  $\xi_1, \xi_2, \dots, \xi_n$  by means of the transformation

$$x = Z\xi \quad (x_i = \sum_{k=1}^n z_{ik}\xi_k, \quad i = 1, 2, \dots, n),$$

where  $Z = \|z_{ik}\|_1^n$  is a principal matrix of the pencil  $A(x, x) - \lambda B(x, x)$ . In the new variables the ratio of the forms is represented (see (63)) by

$$\frac{A(x, x)}{B(x, x)} = \frac{\lambda_1 \xi_1^2 + \lambda_2 \xi_2^2 + \dots + \lambda_n \xi_n^2}{\xi_1^2 + \xi_2^2 + \dots + \xi_n^2}. \quad (70)$$

On the real axis we take the  $n$  points  $\lambda_1, \lambda_2, \dots, \lambda_n$ . We ascribe to these points non-negative masses  $m_1 = \xi_1^2, m_2 = \xi_2^2, \dots, m_n = \xi_n^2$ , respectively. Then, by (70), the quotient  $\frac{A(x, x)}{B(x, x)}$  is the coordinate of the center of these masses. Therefore

$$\lambda_1 \leq \frac{A(x, x)}{B(x, x)} \leq \lambda_n.$$

Let us, for the time being, ignore the second part of the inequality and investigate when the equality sign holds in the first part. For this purpose, we group together the equal characteristic values in (69):

$$\lambda_1 = \dots = \lambda_{p_1} < \lambda_{p_1+1} = \dots = \lambda_{p_1+p_2} < \dots. \quad (71)$$

The center of mass can coincide with the least value  $\lambda_1$  only if all the masses are zero except at this point, i.e., when

$$\xi_{p_1+1} = \dots = \xi_n = 0.$$

In this case the corresponding  $x$  is a linear combination of the principal columns  $z^1, z^2, \dots, z^{p_1}$ .<sup>25</sup> Therefore all these columns correspond to the characteristic value  $\lambda_1$ , so that  $x$  is also a principal column (vector) for  $\lambda = \lambda_1$ .

<sup>24</sup> Here we use the term 'principal vector' in the sense of a principal column of the pencil (see p. 310). Throughout this section, having the geometric interpretation in mind, we often call a column, a vector.

<sup>25</sup> From  $x = Z\xi$  it follows that  $x = \sum_{k=1}^n \xi_k z^k$ .

We have proved:

**THEOREM 10:** *The smallest characteristic value of the regular pencil  $A(x, x) - \lambda B(x, x)$  is the minimum of the ratio of the forms  $A(x, x)$  and  $B(x, x)$*

$$\lambda_1 = \min \frac{A(x, x)}{B(x, x)}. \quad (72)$$

and this minimum is only assumed for principal vectors of the characteristic value  $\lambda_1$ .

**2.** In order to give an analogous 'minimal' characteristic for the next characteristic value  $\lambda_2$ , we restrict ourselves to all the vectors orthogonal to  $z^1$ , i.e., to those that satisfy the equation<sup>26</sup>

$$B(z^1, x) = 0.$$

For these vectors,

$$\frac{A(x, x)}{B(x, x)} = \frac{\lambda_2 \xi_2^2 + \dots + \lambda_n \xi_n^2}{\xi_2^2 + \dots + \xi_n^2}$$

and therefore

$$\min \frac{A(x, x)}{B(x, x)} = \lambda_2 \quad (B(z^1, x) = 0).$$

Here the equality sign holds only for those vectors orthogonal to  $z^1$  that are principal vectors for the characteristic value  $\lambda_2$ .

Proceeding to the subsequent characteristic values, we eventually obtain the following theorem:

**THEOREM 11:** *For every  $p$  ( $1 \leq p \leq n$ ) the  $p$ -th characteristic value  $\lambda_p$  in (69) is the minimum of the ratio of the forms*

$$\lambda_p = \min \frac{A(x, x)}{B(x, x)}, \quad (73)$$

*provided that the variable vector  $x$  is orthogonal to the first  $p - 1$  orthonormal principal vectors  $z^1, z^2, \dots, z^{p-1}$ :*

<sup>26</sup> Here, and in what follows, we shall mean by the orthogonality of two vectors (columns)  $x, y$  that the equation  $B(x, y) = 0$  holds. This is in complete agreement with the geometric interpretation given in the preceding section. We shall regard the quantities  $x_1, x_2, \dots, x_n$  as the coordinates of a vector  $x$  in some basis of a euclidean space in which the square of the length (the norm) is given by the positive-definite form

$B(x, x) = \sum_{i,k=1}^n b_{ik} x_i x_k$ . In this metric the vectors  $z^1, z^2, \dots, z^n$  form an orthonormal

basis. Therefore, if the vector  $x = \sum_{k=1}^n \xi_k z^k$  is orthogonal to one of the  $z^k$ , then the corresponding  $\xi_k = 0$ .

$$B(z^1, x) = 0, \dots, B(z^{p-1}, x) = 0. \quad (74)$$

Moreover, the minimum is assumed only for those vectors that satisfy the condition (74) and are at the same time principal vectors for the characteristic value  $\lambda_p$ .

3. The characterization of  $\lambda_p$  given in Theorem 11 has the disadvantage that it is connected with the preceding principal vectors  $z^1, z^2, \dots, z^{p-1}$  and can therefore be used only when these vectors are known. Moreover, there is a certain arbitrariness in the choice of these vectors.

In order to give a characterization of  $\lambda_p$  ( $p = 1, 2, \dots, n$ ) free from these defects, we introduce the concept of *constraint* imposed on the variables  $x_1, x_2, \dots, x_n$ .

Suppose that linear forms in the variables  $x_1, x_2, \dots, x_n$  are given:

$$L_k(x) = l_{1k}x_1 + l_{2k}x_2 + \dots + l_{nk}x_n \quad (k = 1, 2, \dots, h). \quad (74')$$

We shall say that the variables  $x_1, x_2, \dots, x_n$  or (what is the same) the vector  $x$  is subject to  $h$  constraints  $L_1, L_2, \dots, L_h$  if only such values of the variables are considered that satisfy the system of equations

$$L_k(x) = 0 \quad (k = 1, 2, \dots, h). \quad (74'')$$

Preserving the notation (74') for arbitrary linear forms we introduce a specialized notation for the 'scalar product' of  $x$  with the principal vectors  $z^1, z^2, \dots, z^n$ :

$$\tilde{L}_k(x) = B(z^k, x) \quad (k = 1, 2, \dots, n).^{27} \quad (75)$$

Furthermore, when the variable vector is subject to the constraints (74'') we shall denote  $\min \frac{A(x, x)}{B(x, x)}$  as follows:

$$\mu\left(\frac{A}{B}; L_1, L_2, \dots, L_h\right).$$

In this notation, (73) is written as follows:

$$\lambda_p = \mu\left(\frac{A}{B}; \tilde{L}_1, \tilde{L}_2, \dots, \tilde{L}_{p-1}\right) \quad (p = 1, 2, \dots, n). \quad (76)$$

We consider the constraints

$$L_1(x) = 0, \dots, L_{p-1}(x) = 0 \quad (77)$$

and

$$\tilde{L}_{p+1}(x) = 0, \dots, \tilde{L}_n(x) = 0. \quad (78)$$

<sup>27</sup>  $\tilde{L}_k(x) = z^{kT} Bx = \tilde{l}_{1k}x_1 + \tilde{l}_{2k}x_2 + \dots + \tilde{l}_{nk}x_n$ , where  $\tilde{l}_{1k}, \tilde{l}_{2k}, \dots, \tilde{l}_{nk}$  are the elements of the row matrix  $z^{kT} B$  ( $k = 1, 2, \dots, n$ ).

Since the number of constraints (77) and (78) is less than  $n$ , there exists a vector  $x^{(1)} \neq 0$  satisfying all these constraints. Since the constraints (78) express the orthogonality of  $x$  to the principal vectors  $z^{p+1}, \dots, z^n$ , the corresponding coordinates of  $x^{(1)}$  are  $\xi_{p+1} = \dots = \xi_n = 0$ . Therefore, by (70),

$$\frac{A(x^{(1)}, x^{(1)})}{B(x^{(1)}, x^{(1)})} = \frac{\lambda_1 \xi_1^2 + \dots + \lambda_p \xi_p^2}{\xi_1^2 + \dots + \xi_p^2} \leq \lambda_p.$$

But then

$$\mu\left(\frac{A}{B}; L_1, L_2, \dots, L_{p-1}\right) \leq \frac{A(x^{(1)}, x^{(1)})}{B(x^{(1)}, x^{(1)})} \leq \lambda_p.$$

This inequality in conjunction with (76) shows that for variable constraints  $L_1, L_2, \dots, L_{p-1}$  the value of  $\mu$  remains less than or equal to  $\lambda_p$  and becomes  $\lambda_p$  if the specialized constraints  $\tilde{L}_1, \tilde{L}_2, \dots, \tilde{L}_{p-1}$  are taken.

Thus we have proved:

**THEOREM 12:** *If we consider the minimum of the ratio of the two forms  $\frac{A(x, x)}{B(x, x)}$  for  $p-1$  arbitrary, but variable, constraints  $L_1, L_2, \dots, L_{p-1}$ , then the maximum of these minima is equal to  $\lambda_p$ :*

$$\lambda_p = \max \mu\left(\frac{A}{B}; L_1, L_2, \dots, L_{p-1}\right) \quad (p = 1, \dots, n). \quad (79)$$

Theorem 12 gives a 'maximal-minimal' characterization of  $\lambda_1, \lambda_2, \dots, \lambda_n$  in contrast to the 'minimal' characterization which we discussed in Theorem 11.

4. Note that when in the pencil  $A(x, x) - \lambda B(x, x)$  the form  $A(x, x)$  is replaced by  $-A(x, x)$ , all the characteristic values of the pencil change sign, but the corresponding principal vectors remain unchanged. Thus, the characteristic values of the pencil  $-A(x, x) - \lambda B(x, x)$  are

$$-\lambda_n \leq -\lambda_{n-1} \leq \dots \leq -\lambda_1.$$

Moreover, by using the notation

$$v\left(\frac{A}{B}; L_1, L_2, \dots, L_h\right) = \max \frac{A(x, x)}{B(x, x)} \quad (80)$$

when the variable vector is subject to the constraints  $L_1, L_2, \dots, L_h$ , we can write:

$$\mu\left(-\frac{A}{B}; L_1, L_2, \dots, L_h\right) = -v\left(\frac{A}{B}; L_1, L_2, \dots, L_h\right)$$

and

$$\max \mu\left(-\frac{A}{B}; L_1, L_2, \dots, L_h\right) = -\min v\left(\frac{A}{B}; L_1, L_2, \dots, L_h\right).$$

Therefore, by applying Theorems 10, 11, and 12 to the ratio  $-\frac{A(x, x)}{B(x, x)}$  we obtain instead of (72), (76), and (79) the formulas

$$\lambda_n = \max \frac{A(x, x)}{B(x, x)},$$

$$\lambda_{n-p+1} = \nu \left( \frac{A}{B}; \tilde{L}_n, \tilde{L}_{n-1}, \dots, \tilde{L}_{n-p+2} \right) \quad (p = 2, \dots, n).$$

$$\lambda_{n-p+1} = \min \nu \left( \frac{A}{B}; L_1, L_2, \dots, L_{p-1} \right),$$

These formulas establish the 'maximal' and the 'minimal-maximal' properties, respectively, of  $\lambda_1, \lambda_2, \dots, \lambda_n$ , which we formulate in the following theorem:

THEOREM 13: *Suppose that to the characteristic values*

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$$

*of the regular pencil of forms  $A(x, x) - \lambda B(x, x)$  there correspond the linearly independent principal vectors of the pencil  $z^1, z^2, \dots, z^n$ . Then:*

1) *The largest characteristic value  $\lambda_n$  is the maximum of the ratio of the forms  $\frac{A(x, x)}{B(x, x)}$ :*

$$\lambda_n = \max \frac{A(x, x)}{B(x, x)}, \quad (81)$$

*and this maximum is assumed only for principal vectors of the pencil corresponding to the characteristic value  $\lambda_n$ .*

2) *The characteristic value  $p$ -th from the end  $\lambda_{n-p+1}$  ( $2 \leq p \leq n$ ) is the maximum of the same ratio of the forms*

$$\lambda_{n-p+1} = \max \frac{A(x, x)}{B(x, x)} \quad (82)$$

*provided that the variable vector  $x$  is subject to the constraints:*<sup>28</sup>

$$B(z^n, x) = 0, B(z^{n-1}, x) = 0, \dots, B(z^{n-p+2}, x) = 0, \quad (83)$$

*i.e.,*

$$\lambda_{n-p+1} = \nu \left( \frac{A}{B}; \tilde{L}_n, \tilde{L}_{n-1}, \dots, \tilde{L}_{n-p+2} \right); \quad (84)$$

*this maximum is assumed only for principal vectors of the pencil corresponding to the characteristic value  $\lambda_{n-p+1}$  and satisfying the constraints (83).*

<sup>28</sup> In a euclidean space with a metric form  $B(x, x)$ , the condition (83) expresses the fact that the vector  $x$  is orthogonal to the principal vectors  $z^{n-p+2}, \dots, z^n$ . See footnote 26.

3) If in the maximum of the ratio of the forms  $\frac{A(x, x)}{B(x, x)}$  with the constraints

$$L_1(x) = 0, \dots, L_{p-1}(x) = 0 \quad (2 \leq p \leq n)$$

( $2 \leq p \leq n$ ) the constraints are varied, then the least value (minimum) of this maximum is equal to  $\lambda_{n-p+1}$ :

$$\lambda_{n-p+1} = \min \nu \left( \frac{A}{B}; L_1, L_2, \dots, L_{p-1} \right). \quad (85)$$

5. Let

$$L_1^0(x) = 0, L_2^0(x) = 0, \dots, L_h^0(x) = 0. \quad (86)$$

be  $h$  independent constraints.<sup>29</sup> Then we can express  $h$  of the variables  $x_1, x_2, \dots, x_n$  by the remaining variables, which we denote by  $v_1, v_2, \dots, v_{n-h}$ . Therefore, when the constraints (86) are imposed, the regular pencil of forms  $A(x, x) - \lambda B(x, x)$  goes over into the pencil  $A^0(v, v) - \lambda B^0(v, v)$ , where  $B^0(v, v)$  is again a positive-definite form (only in  $n - h$  variables). The regular pencil so obtained has  $n - h$  real characteristic values

$$\lambda_1^0 \leq \lambda_2^0 \leq \dots \leq \lambda_{n-h}^0. \quad (87)$$

Subject to the constraints (86) we can express all the variables in terms of  $n - h$  independent ones  $v_1, v_2, \dots, v_{n-h}$  in various ways. However, the characteristic values (87) are independent of this arbitrariness and have completely definite values. This follows, for example, from the maximal-minimal property of the characteristic values

$$\lambda_1^0 = \min \frac{A^0(v, v)}{B^0(v, v)} = \mu \left( \frac{A}{B}; L_1^0, L_2^0, \dots, L_h^0 \right) \quad (88)$$

and, in general,

$$\begin{aligned} \lambda_p^0 &= \max \mu \left( \frac{A^0}{B^0}; L_1, L_2, \dots, L_{p-1} \right) \\ &= \max \mu \left( \frac{A}{B}; L_1^0, \dots, L_h^0, L_1, \dots, L_{p-1} \right), \end{aligned} \quad (89)$$

where in (89) only the constraints  $L_1, L_2, \dots, L_{p-1}$  are allowed to vary.

<sup>29</sup> The constraints (86) are independent when the linear forms  $L_1^0(x), L_2^0(x), \dots, L_h^0(x)$  on the left-hand sides of (86) are independent.

The following theorem holds:

**THEOREM 14:** *If  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  are the characteristic values of the regular pencil of forms  $A(x, x) - \lambda B(x, x)$  and  $\lambda_1^0 \leq \lambda_2^0 \leq \dots \leq \lambda_{n-h}^0$  are the characteristic values of the same pencil subject to  $h$  independent constraints, then*

$$\lambda_p \leq \lambda_p^0 \leq \lambda_{p+h} \quad (p=1, 2, \dots, n-h). \quad (90)$$

*Proof.* The inequality  $\lambda_p \leq \lambda_p^0$  ( $p=1, 2, \dots, n-h$ ) follows easily from (79) and (89). For when new constraints are added, the value of the minimum  $\mu\left(\frac{A}{B}; L_1, \dots, L_{p-1}\right)$  increases or remains the same. Therefore

$$\mu\left(\frac{A}{B}; L_1, \dots, L_{p-1}\right) \leq \mu\left(\frac{A}{B}; L_1^0, \dots, L_h^0; L_1, \dots, L_{p-1}\right).$$

Hence

$$\lambda_p = \max \mu\left(\frac{A}{B}; L_1, \dots, L_{p-1}\right) \leq \lambda_p^0 = \max \mu\left(\frac{A}{B}; L_1^0, \dots, L_h^0; L_1, \dots, L_{p-1}\right).$$

The second part of the inequality (90) holds in view of the relations

$$\begin{aligned} \lambda_p^0 &= \max \mu\left(\frac{A}{B}; L_1^0, \dots, L_h^0; L_1, \dots, L_{p-1}\right) \\ &\leq \max \mu\left(\frac{A}{B}; L_1, \dots, L_{p-1}, L_p, \dots, L_{p+h-1}\right) = \lambda_{p+h}. \end{aligned}$$

Here not only are  $L_1, \dots, L_{p-1}$  varied, on the right-hand side, but  $L_p, \dots, L_{p+h-1}$  also; on the left-hand side the latter are replaced by the fixed constraints  $L_1^0, L_2^0, \dots, L_h^0$ .

This completes the proof.

6. Suppose that two regular pencils of forms

$$A(x, x) - \lambda B(x, x), \quad \tilde{A}(x, x) - \lambda \tilde{B}(x, x) \quad (91)$$

are given and that for every  $x \neq 0$ ,

$$\frac{A(x, x)}{B(x, x)} \leq \frac{\tilde{A}(x, x)}{\tilde{B}(x, x)}.$$

Then obviously,

$$\max \mu \left( \frac{A}{B}; L_1, L_2, \dots, L_{p-1} \right) \leq \max \mu \left( \frac{\tilde{A}}{\tilde{B}}; L_1, L_2, \dots, L_{p-1} \right) \\ (p = 1, 2, \dots, n).$$

Therefore, if we denote by  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  and  $\tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \dots \leq \tilde{\lambda}_n$ , respectively, the characteristic values of the pencils (91), then we have:

$$\lambda_p \leq \tilde{\lambda}_p \quad (p = 1, 2, \dots, n).$$

Thus, we have proved the following theorem:

**THEOREM 15:** *If two regular pencils of forms  $A(x, x) - \lambda B(x, x)$  and  $\tilde{A}(x, x) - \lambda \tilde{B}(x, x)$  with the characteristic values  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  and  $\tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \dots \leq \tilde{\lambda}_n$  are given, then the identical relation*

$$\frac{A(x, x)}{B(x, x)} \leq \frac{\tilde{A}(x, x)}{\tilde{B}(x, x)} \quad (92)$$

*implies that*

$$\lambda_p \leq \tilde{\lambda}_p \quad (p = 1, 2, \dots, n). \quad (93)$$

Let us consider the special case where, in (92),  $B(x, x) \equiv \tilde{B}(x, x)$ . In this case, the difference  $\tilde{A}(x, x) - A(x, x)$  is a positive-semidefinite quadratic form and can therefore be expressed as a sum of independent positive squares:

$$\tilde{A}(x, x) = A(x, x) + \sum_{i=1}^r [X_i(x)]^2.$$

Then, when the  $r$  independent constraints

$$X_1(x) = 0, X_2(x) = 0, \dots, X_r(x) = 0$$

are imposed, the forms  $A(x, x)$  and  $\tilde{A}(x, x)$  coincide, and the pencils  $A(x, x) - \lambda B(x, x)$  and  $\tilde{A}(x, x) - \lambda B(x, x)$  have the same characteristic values

$$\lambda_1^0 \leq \lambda_2^0 \leq \dots \leq \lambda_{n-r}^0.$$

Applying Theorem 14 to both pencils  $A(x, x) - \lambda B(x, x)$  and  $\tilde{A}(x, x) - \lambda B(x, x)$ , we have:

$$\tilde{\lambda}_p \leq \lambda_p^0 \leq \lambda_{p+r} \quad (p = 1, 2, \dots, n-r).$$

In conjunction with the inequality (93), this leads to the following theorem:



**THEOREM 16:** If  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  and  $\tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \dots \leq \tilde{\lambda}_n$  are the characteristic values of two regular pencils of forms  $A(x, x) - \lambda B(x, x)$  and  $\tilde{A}(x, x) - \lambda B(x, x)$ , where

$$\tilde{A}(x, x) = A(x, x) + \sum_{i=1}^r [X_i(x)]^2,$$

and  $X_i(x)$  ( $i = 1, 2, \dots, r$ ) are independent linear forms, then the following inequalities hold:<sup>30</sup>

$$\lambda_p \leq \tilde{\lambda}_p \leq \lambda_{p+r} \quad (p = 1, 2, \dots, n). \quad (94)$$

In exactly the same way the following theorem is proved:

**THEOREM 17:** If  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  and  $\tilde{\lambda}_1 \leq \tilde{\lambda}_2 \leq \dots \leq \tilde{\lambda}_n$  are the characteristic values of the regular pencil of forms  $A(x, x) - \lambda B(x, x)$  and  $A(x, x) - \lambda \tilde{B}(x, x)$ , where the form  $\tilde{B}(x, x)$  is obtained from  $B(x, x)$  by adding  $r$  positive squares, then the following inequalities hold:<sup>31</sup>

$$\lambda_{p-r} \leq \tilde{\lambda}_p \leq \lambda_p \quad (p = 1, 2, \dots, n). \quad (95)$$

*Note.* In Theorems 16 and 17 we can claim that for some  $p$  we have, respectively  $\lambda_p < \tilde{\lambda}_p$  and  $\tilde{\lambda}_p < \lambda_p$ , provided of course that  $r \neq 0$ .<sup>32</sup>

## § 8. Small Oscillations of a System with $n$ Degrees of Freedom

The results of the two preceding sections have important applications in the theory of small oscillations of a mechanical system with  $n$  degrees of freedom.

**1.** We consider the free oscillations of a conservative mechanical system with  $n$  degrees of freedom near a stable position of equilibrium. We shall give the deviation of the system from the position of equilibrium by means of independent generalized coordinates  $q_1, q_2, \dots, q_n$ . The position of equilibrium itself corresponds to zero values of these coordinates:  $q_1 = 0, q_2 = 0, \dots, q_n = 0$ . Then the kinetic energy of the system is represented as a quadratic form in the generalized velocities  $\dot{q}_1, \dot{q}_2, \dots, \dot{q}_n$ :<sup>33</sup>

$$T = \sum_{i,k=1}^n b_{ik}(q_1, q_2, \dots, q_n) \dot{q}_i \dot{q}_k.$$

<sup>30</sup> The second parts of these inequalities hold for  $p \leq n - r$  only.

<sup>31</sup> The first parts of the inequalities hold for  $p > r$ .

<sup>32</sup> See [17], pp. 71-73.

<sup>33</sup> A dot denotes the derivative with respect to time.

Expanding the coefficients  $b_{ik}(q_1, q_2, \dots, q_n)$  as power series in  $q_1, q_2, \dots, q_n$

$$b_{ik}(q_1, q_2, \dots, q_n) = b_{ik} + \dots \quad (i, k = 1, 2, \dots, n)$$

and keeping only the constant terms  $b_{ik}$ , since the deviations  $q_1, q_2, \dots, q_n$  are small, we then have:

$$T = \sum_{i,k=1}^n b_{ik} \dot{q}_i \dot{q}_k \quad (b_{ik} = b_{ki}; i, k = 1, 2, \dots, n).$$

The kinetic energy is always positive, and is zero only for zero velocities  $\dot{q}_1 = \dot{q}_2 = \dots = \dot{q}_n = 0$ . Therefore  $\sum_{i,k=1}^n b_{ik} \dot{q}_i \dot{q}_k$  is a positive-definite form.

The potential energy of the system is a function of the coordinates:  $P(q_1, q_2, \dots, q_n)$ . Without loss of generality, we can take

$$P_0 = P(0, 0, \dots, 0) = 0.$$

Then, expanding the potential energy as a power series in  $q_1, q_2, \dots, q_n$ , we obtain:

$$P = \sum_{i=1}^n a_i q_i + \sum_{i,k=1}^n a_{ik} q_i q_k + \dots$$

Since in a position of equilibrium the potential energy always has a stationary value, we have

$$a_i = \left. \frac{\partial P}{\partial q_i} \right|_0 = 0 \quad (i = 1, 2, \dots, n).$$

Keeping only the terms of the second order in  $q_1, q_2, \dots, q_n$ , we have

$$P = \sum_{i,k=1}^n a_{ik} q_i q_k \quad (a_{ik} = a_{ki}; i, k = 1, 2, \dots, n).$$

Thus, the potential energy  $P$  and the kinetic energy  $T$  are determined by two quadratic forms:

$$P = \sum_{i,k=1}^n a_{ik} q_i q_k, \quad T = \sum_{i,k=1}^n b_{ik} \dot{q}_i \dot{q}_k, \quad (96)$$

the second of which is positive definite.

We now write down the differential equations of motion in the form of Lagrange's equations of the second kind:<sup>34</sup>

$$\frac{d}{dt} \frac{\partial T}{\partial \dot{q}_i} - \frac{\partial T}{\partial q_i} = - \frac{\partial P}{\partial q_i} \quad (i = 1, 2, \dots, n). \quad (97)$$

<sup>34</sup> See, for example, G. K. Suslov (Suslov), *Theoretische Mechanik*, § 191.

When we substitute for  $T$  and  $P$  their expressions from (96), we obtain:

$$\sum_{k=1}^n b_{ik} \ddot{q}_k + \sum_{k=1}^n a_{ik} q_k = 0 \quad (i=1, 2, \dots, n). \quad (98)$$

We introduce the real symmetric matrices

$$A = \| a_{ik} \|_1^n \text{ and } B = \| b_{ik} \|_1^n$$

and the column matrix  $q = (q_1, q_2, \dots, q_n)$  and write the system of equations (98) in the following matrix form:

$$B\ddot{q} + Aq = 0. \quad (98')$$

We shall seek solutions of (98) in the form of harmonic oscillations

$$q_1 = v_1 \sin(\omega t + \alpha), \quad q_2 = v_2 \sin(\omega t + \alpha), \quad \dots \quad q_n = v_n \sin(\omega t + \alpha),$$

in matrix notation:

$$q = v \sin(\omega t + \alpha). \quad (99)$$

Here  $v = (v_1, v_2, \dots, v_n)$  is the constant-amplitude column (constant-amplitude 'vector'),  $\omega$  is the frequency, and  $\alpha$  is the initial phase of the oscillation.

Substituting the expression (99) for  $q$  in (98') and cancelling  $\sin(\omega t + \alpha)$ , we obtain:

$$Av = \lambda Bv \quad (\lambda = \omega^2).$$

But this equation is the same as (49). Therefore the required amplitude vector is a principal vector, and the square of the frequency  $\lambda = \omega^2$  is the corresponding characteristic value of the regular pencil of forms  $A(x, x) - \lambda B(x, x)$ .

We subject the potential energy to an additional restriction by postulating that the function  $P(q_1, q_2, \dots, q_n)$  in a position of equilibrium shall have a strict minimum.<sup>35</sup>

Then, by a theorem of Dirichlet,<sup>36</sup> the position of equilibrium is stable. On the other hand, our assumption means that the quadratic form  $P = A(q, q)$  is also positive definite.

By Theorem 8, the regular pencil of forms  $A(x, x) - \lambda B(x, x)$  has real characteristic values  $\lambda_1, \lambda_2, \dots, \lambda_n$  and  $n$  corresponding principal characteristic vectors  $v^1, v^2, \dots, v^n$  ( $v^k = (v_{1k}, v_{2k}, \dots, v_{nk})$ ;  $k = 1, 2, \dots, n$ ) satisfying the condition

<sup>35</sup> I.e., that the value of  $P_0$  in the position of equilibrium is less than all other values of the function in some neighborhood of the position of equilibrium.

<sup>36</sup> See G. K. Suslov (Suslov), *Theoretische Mechanik*, § 210.

$$B(v^i, v^k) = \sum_{\mu, \nu=1}^n b_{\mu\nu} v_{\mu i} v_{\nu k} = \delta_{ik} \quad (i, k = 1, 2, \dots, n). \quad (100)$$

From the fact that  $A(x, x)$  is positive definite it follows that all the characteristic values of the pencil  $A(x, x) - \lambda B(x, x)$  are positive.<sup>37</sup>

$$\lambda_k > 0 \quad (k = 1, 2, \dots, n).$$

But then there exist  $n$  harmonic oscillations<sup>38</sup>

$$v^k \sin(\omega_k t + \alpha_k) \quad (\omega_k^2 = \lambda_k, k = 1, 2, \dots, n), \quad (101)$$

whose amplitude vectors  $v^k = (v_{1k}, v_{2k}, \dots, v_{nk})$  ( $k = 1, 2, \dots, n$ ) satisfy the conditions of 'orthonormality' (100).

Since the equation (98') is linear, every oscillation can be obtained by a superposition of the harmonic oscillations (101):

$$q = \sum_{k=1}^n A_k \sin(\omega_k t + \alpha_k) v^k, \quad (102)$$

where  $A_k$  and  $\alpha_k$  are arbitrary constants. For, whatever the values of these constants, the expression (102) is a solution of (98'). On the other hand, the arbitrary constants can be made to satisfy the following initial conditions:

$$q|_{t=0} = q_0, \quad \dot{q}|_{t=0} = \dot{q}_0.$$

For from (102) we find:

$$q_0 = \sum_{k=1}^n A_k \sin \alpha_k v^k, \quad \dot{q}_0 = \sum_{k=1}^n \omega_k A_k \cos \alpha_k v^k. \quad (103)$$

Since the principal columns  $v^1, v^2, \dots, v^n$  are always linearly independent, the values  $A_k \sin \alpha_k$  and  $\omega_k \cos \alpha_k$  ( $k = 1, 2, \dots, n$ ), and hence the constants  $A_k$  and  $\alpha_k$  ( $k = 1, 2, \dots, n$ ), are uniquely determined from (103).

The solution (102) of our system of differential equations can be written more conveniently:

$$q_i = \sum_{k=1}^n A_k \sin(\omega_k t + \alpha_k) v_{ik}. \quad (104)$$

Note that we could also derive the formulas (102) and (104) starting from Theorem 9. For let us consider a non-singular transformation of the

<sup>37</sup> This follows, for example, from the representation (63).

<sup>38</sup> Here the initial phases  $\alpha_k$  ( $k = 1, 2, \dots, n$ ) are arbitrary constants.

variables with the matrix  $V = \|v_{ik}\|_1^n$  that reduces the two forms  $A(x, x)$  and  $B(x, x)$  simultaneously to the canonical form (63). Setting

$$q_i = \sum_{k=1}^n v_{ik} \theta_k \quad (i=1, 2, \dots, n) \quad (105)$$

or, more briefly,

$$q = V\theta \quad (\theta = (\theta_1, \theta_2, \dots, \theta_n)) \quad (106)$$

and observing that  $\dot{q} = V\dot{\theta}$ , we have:

$$P = A(q, q) = \sum_{k=1}^n \lambda_k \theta_k^2, \quad T = B(\dot{q}, \dot{q}) = \sum_{k=1}^n \dot{\theta}_k^2. \quad (107)$$

The coordinates  $\theta_1, \theta_2, \dots, \theta_n$  in which the potential and kinetic energies have a representation as in (107) are called *principal coordinates*.

We now make use of Lagrange's equations of the second kind (98) and substitute the expressions (107) for  $P$  and  $T$ . We obtain:

$$\ddot{\theta}_k + \lambda_k \theta_k = 0 \quad (k=1, 2, \dots, n). \quad (108)$$

Since  $A(q, q)$  is positive definite, all the numbers  $\lambda_1, \lambda_2, \dots, \lambda_n$  are positive and can be represented in the form

$$\lambda_k = \omega_k^2 \quad (\omega_k > 0; k=1, 2, \dots, n). \quad (109)$$

From (108) and (109), we find:

$$\theta_k = A_k \sin(\omega_k t + \alpha_k) \quad (k=1, 2, \dots, n). \quad (110)$$

When we substitute these expressions for  $\theta_k$  in (105), we again obtain the formulas (104) and therefore (102). The values  $v_{ik}$  ( $i, k=1, 2, \dots, n$ ) in both methods are the same, because the matrix  $V = \|v_{ik}\|_1^n$  in (106) is, by Theorem 9, a principal matrix of the regular pencil of forms  $A(x, x) - \lambda B(x, x)$ .

2. We also mention a mechanical interpretation of Theorems 14 and 15.

We number the frequencies  $\omega_1, \omega_2, \dots, \omega_n$  of the given mechanical system in non-descending order:

$$0 < \omega_1 \leq \omega_2 \leq \dots \leq \omega_n.$$

The disposition of the corresponding characteristic values  $\lambda_k = \omega_k^2$  ( $k=1, 2, 3, \dots, n$ ) of the pencil  $A(x, x) - \lambda B(x, x)$  is then also determined:

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n.$$

We impose  $h$  independent finite stationary constraints<sup>39</sup> on the given system. Since the deviations  $q_1, q_2, \dots, q_n$  are supposed to be small, these connections can be assumed to be linear in  $q_1, q_2, \dots, q_n$ :

$$L_1(q) = 0, L_2(q) = 0, \dots, L_h(q) = 0.$$

After the constraints are imposed, our system has  $n - h$  degrees of freedom. The frequencies of the system,

$$\omega_1^0 \leq \omega_2^0 \leq \dots \leq \omega_{n-h}^0,$$

are connected with the characteristic values  $\lambda_1^0 \leq \lambda_2^0 \leq \dots \leq \lambda_{n-h}^0$  of the pencil  $A(x, x) - \lambda B(x, x)$ , subject to the constraints  $L_1, L_2, \dots, L_h$ , by the relations  $\lambda_j^0 = \omega_j^{02}$  ( $j = 1, 2, \dots, n - h$ ). Therefore Theorem 14 immediately implies that

$$\omega_j \leq \omega_j^0 \leq \omega_{j+h} \quad (j = 1, 2, \dots, n - h).$$

Thus: *When  $h$  constraints are imposed, the frequencies of a system can only increase, but the value of the new  $j$ -th frequency  $\omega_j^0$  cannot exceed the value of the previous  $(j + h)$ -th frequency  $\omega_{j+h}$ .*

In exactly the same way, we can assert on the basis of Theorem 15 that: *With increasing rigidity of the system, i.e., with an increase of the form  $A(q, q)$  for the potential energy (without a change in  $B(\dot{q}, \dot{q})$ ), the frequencies can only increase; and with increasing inertia of the system, i.e., with an increase of the form  $B(\dot{q}, \dot{q})$  for the kinetic energy (without a change in  $A(q, q)$ ), the frequencies can only decrease.*

Theorems 16 and 17 lead to an additional sharpening of this proposition.<sup>40</sup>

## § 9. Hermitian Forms<sup>41</sup>

1. All the results of §§ 1-7 of this chapter that were established for quadratic forms can be extended to hermitian forms.

We recall<sup>42</sup> that a *hermitian form* is an expression

<sup>39</sup> A finite stationary constraint is expressed by an equation  $f(q_1, q_2, \dots, q_n) = 0$ , where  $f(q_1, q_2, \dots, q_n)$  is some function of the generalized coordinates.

<sup>40</sup> The reader can find an account of the oscillatory properties of elastic oscillations of a system with  $n$  degrees of freedom in [17], Chapter III.

<sup>41</sup> In the preceding sections, all the numbers and variables were real. In this section, the numbers are complex and the variables assume complex values.

<sup>42</sup> See Chapter IX, § 2.

$$H(x, x) = \sum_{i, k=1}^n h_{ik} x_i \bar{x}_k \quad (h_{ik} = \bar{h}_{ki}; \quad i, k = 1, 2, \dots, n). \quad (111)$$

To the hermitian form (111) there corresponds the following *bilinear hermitian form*:

$$H(x, y) = \sum_{i, k=1}^n h_{ik} x_i \bar{y}_k; \quad (112)$$

moreover,

$$H(y, x) = \overline{H(x, y)} \quad (113)$$

and, in particular,

$$H(x, x) = \overline{H(x, x)} \quad (113')$$

i.e., the hermitian form  $H(x, x)$  assumes real values only.

The coefficient matrix  $H = \| h_{ik} \|_1^n$  of the hermitian form is hermitian, i.e.,  $H^* = H$ .<sup>43</sup>

By means of the matrix  $H = \| h_{ik} \|_1^n$  we can represent  $H(x, y)$  and, in particular,  $H(x, x)$  in the form of a product of three matrices, a row, a square, and a column matrix:<sup>44</sup>

$$H(x, y) = x^T H \bar{y}, \quad H(x, x) = x^T H \bar{x}. \quad (114)$$

If

$$x = \sum_{i=1}^m c_i u^i, \quad y = \sum_{k=1}^p d_k v^k, \quad (115)$$

where  $u^i, v^k$  are column matrices and  $c_i, d_k$  are complex numbers ( $i = 1, 2, 3, \dots, m; k = 1, 2, \dots, p$ ), then

$$H(x, y) = \sum_{i=1}^m \sum_{k=1}^p c_i \bar{d}_k H(u^i, v^k). \quad (116)$$

We subject the variables  $x_1, x_2, \dots, x_n$  to the linear transformation

$$x_i = \sum_{k=1}^n t_{ik} \bar{x}_k \quad (i = 1, 2, \dots, n) \quad (117)$$

<sup>43</sup> A matrix symbol followed by an asterisk \* denotes the matrix that is obtained from the given one by transposition and replacement of all the elements by their complex conjugates ( $H^* = \bar{H}^T$ ).

<sup>44</sup> Here

$x = (x_1, x_2, \dots, x_n), \quad \bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n), \quad y = (y_1, y_2, \dots, y_n), \quad \bar{y} = (\bar{y}_1, \bar{y}_2, \dots, \bar{y}_n);$

the sign  $T$  denotes transposition.

or, in matrix notation,

$$x = T\xi \quad (T = \|t_{ik}\|_1^n). \quad (117')$$

After the transformation,  $H(x, x)$  assumes the form

$$\tilde{H}(\xi, \xi) = \sum_{i, k=1}^n \tilde{h}_{ik} \xi_i \bar{\xi}_k,$$

where the new coefficient matrix  $\tilde{H} = \|\tilde{h}_{ik}\|_1^n$  is connected with the old coefficient matrix  $H = \|h_{ik}\|_1^n$  by the formula

$$\tilde{H} = T^* H T. \quad (118)$$

This is immediately clear when, in the second of the formulas (114),  $x$  is replaced by  $T\xi$ .

If we set  $T = \bar{W}$ , then we can rewrite (118) as follows:

$$\tilde{H} = W^* H W. \quad (119)$$

From the formula (118) it follows that  $H$  and  $\tilde{H}$  have the same rank provided the transformation (117) is non-singular ( $|T| \neq 0$ ). The rank of  $H$  is called the *rank of the form*  $H(x, x)$ .

The determinant  $|H|$  is called the *discriminant* of  $H(x, x)$ . From (118) we obtain the formula for the transformation of the discriminant on transition to new variables:

$$|\tilde{H}| = |H| |T| |\bar{T}|.$$

A hermitian form is called *singular* if its discriminant is zero. Obviously, a singular form remains singular under any transformation of the variables (117).

A hermitian form  $H(x, x)$  can be represented in infinitely many ways in the form

$$H(x, x) = \sum_{i=1}^r a_i X_i \bar{X}_i, \quad (120)$$

where  $a_i \neq 0$  ( $i = 1, 2, \dots, r$ ) are real numbers and

$$X_i = \sum_{k=1}^n a_{ik} x_k \quad (i = 1, 2, \dots, r)$$

are independent complex linear forms in the variables  $x_1, x_2, \dots, x_n$ .<sup>45</sup>

<sup>45</sup> Therefore  $r \leq n$ .



We shall call the right-hand side of (120) a *sum of linearly independent squares*<sup>46</sup> and every term in the sum a *positive* or a *negative* square according as  $a_i > 0$  or  $< 0$ . Just as for quadratic forms, the number  $r$  in (120) is equal to the rank of the form  $H(x, x)$ .

THEOREM 18 (The Law of Inertia for Hermitian Forms): *In the representation of a hermitian form  $H(x, x)$  as a sum of linearly independent squares,*

$$H(x, x) = \sum_{i=1}^r a_i \bar{X}_i X_i,$$

*the number of positive squares and the number of negative squares do not depend on the choice of the representation.*

The proof is completely analogous to the proof of Theorem 1 (p. 297).

The difference  $\sigma$  between the number  $\pi$  of positive squares and the number  $\nu$  of negative squares in (120) is called the *signature* of the hermitian form  $H(x, x)$ :  $\sigma = \pi - \nu$ .

Lagrange's method of reduction of quadratic forms to sums of squares can also be used for hermitian forms, only the fundamental formulas (15) and (16) on p. 299 must then be replaced by the formulas<sup>47</sup>

$$H(x, x) = \frac{1}{h_{gg}} \left| \sum_{k=1}^n h_{kg} x_k \right|^2 + H_1(x, x), \quad (121)$$

$$H(x, x) = \frac{1}{2} \left\{ \left| \sum_{k=1}^n \left( h_{kf} + \frac{h_{kg}}{h_{fg}} \right) x_k \right|^2 - \left| \sum_{k=1}^n \left( h_{kf} - \frac{h_{kg}}{h_{fg}} \right) x_k \right|^2 \right\} + H_2(x, x). \quad (122)$$

Let us proceed to establish Jacobi's formula for a hermitian form  $H(x, x) = \sum_{i,k=1}^n h_{ik} x_i \bar{x}_k$  of rank  $r$ . Here, as in the case of a quadratic form, we assume that

$$D_k = H \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \neq 0 \quad (k = 1, 2, \dots, r). \quad (123)$$

This inequality enables us to use Theorem 2 of Chapter II (p. 38) on the representation of an arbitrary square matrix in the form of a product of three matrices: a lower triangular matrix  $F$ , a diagonal matrix  $D$ , and an upper triangular matrix  $L$ . We apply this theorem to the matrix  $H = \|h_{ik}\|_1^r$

<sup>46</sup> This terminology is connected with the fact that  $\bar{X}_i X_i$  is the square of the modulus of  $X_i$  ( $\bar{X}_i X_i = |X_i|^2$ ).

<sup>47</sup> The formula (121) is applicable when  $h_{gg} \neq 0$ ; and (122), when  $h_{ff} = h_{gg} = 0$   $h_{fg} \neq 0$ .

and obtain

$$H = F \left\{ D_1, \frac{D_2}{D_1}, \dots, \frac{D_r}{D_{r-1}}, 0, \dots, 0 \right\} L, \quad (124)$$

where  $F = \| f_{ik} \|_1^n$ ,  $L = \| l_{ik} \|_1^n$ , and

$$f_{jk} = \frac{1}{D_k} H \begin{pmatrix} 1 & \dots & k-1 & j \\ 1 & \dots & k-1 & k \end{pmatrix}, \quad l_{kj} = \frac{1}{D_k} H \begin{pmatrix} 1 & \dots & k-1 & k \\ 1 & \dots & k-1 & j \end{pmatrix} \quad (125)$$

$$(j = k, k+1, \dots, n; k = 1, 2, \dots, r),$$

$$f_{ik} = l_{ki} = 0 \quad (i < k; i, k = 1, 2, \dots, n). \quad (126)$$

Since  $H = \| h_{ik} \|_1^n$  is a hermitian matrix, it follows from these equations that

$$f_{ik} = \bar{l}_{ki} \quad \begin{pmatrix} i \geq k; k = 1, 2, \dots, r; i = 1, 2, \dots, n, \\ i < k; i, k = 1, 2, \dots, n \end{pmatrix}. \quad (127)$$

Since all the elements in the last  $n - r$  columns of  $F$  and the last  $n - r$  rows of  $L$  can be chosen arbitrarily,<sup>48</sup> we choose these elements such that 1) the relations (127) hold for all  $i, k$

$$f_{ik} = \bar{l}_{ki} \quad (i, k = 1, 2, \dots, n)$$

and 2)  $|F| = |L| \neq 0$ . Then

$$F = L^*, \quad (128)$$

and (124) assumes the form

$$H = L^* \left\{ D_1, \frac{D_2}{D_1}, \dots, \frac{D_r}{D_{r-1}}, 0, \dots, 0 \right\} L. \quad (129)$$

Setting

$$T = \| t_{ik} \|_1^n = \bar{L}, \quad (130)$$

we write (129) as follows:

$$H = T^T \left\{ D_1, \frac{D_2}{D_1}, \dots, \frac{D_r}{D_{r-1}}, 0, \dots, 0 \right\} \bar{T} \quad (|T| \neq 0). \quad (131)$$

A comparison of this formula with (118) shows that the hermitian form

$$\sum_{k=1}^n \frac{D_k}{D_{k-1}} \xi_k \bar{\xi}_k \quad (D_0 = 1) \quad (132)$$

under the transformation of the variables

<sup>48</sup> These elements, in fact, drop out of the right-hand side of (124), because the last  $n - r$  diagonal elements of  $D$  are zero.

$$\xi_i = \sum_{k=1}^n t_{ik} x_k \quad (i=1, 2, \dots, n)$$

goes over into  $H(x, x)$ , i.e., that Jacobi's formula holds:

$$H(x, x) = \sum_{k=1}^n \frac{D_k}{D_{k-1}} X_k \bar{X}_k \quad (D_0 = 1), \quad (133)$$

where

$$X_k = x_k + t_{k,k+1} x_{k+1} + \dots + t_{kn} x_n \quad (k=1, 2, \dots, r) \quad (134)$$

and

$$t_{kj} = \frac{1}{D_k} H \begin{pmatrix} 1 & 2 & \dots & k-1 & j \\ 1 & 2 & \dots & k-1 & k \end{pmatrix} \quad (j=k+1, \dots, n; k=1, 2, \dots, r). \quad (135)$$

The linear forms  $X_1, X_2, \dots, X_r$  are independent, since  $X_k$  contains the variable  $x_k$  which does not occur in the subsequent forms  $X_{k+1}, \dots, X_r$ .

When we introduce, in place of  $X_1, X_2, \dots, X_r$ , the linearly independent forms

$$Y_k = D_k X_k \quad (k=1, 2, \dots, r), \quad (136)$$

we can write Jacobi's formula (133) in the form

$$H(x, x) = \sum_{k=1}^r \frac{Y_k \bar{Y}_k}{D_{k-1} D_k} \quad (D_0 = 1). \quad (137)$$

According to Jacobi's formula (137), the number of negative squares in the representation of  $H(x, x)$  is equal to the number of variations of sign in the sequence  $1, D_1, D_2, \dots, D_r$ ,

$$\nu = \mathfrak{K}(1, D_1, D_2, \dots, D_r),$$

so that the signature  $\sigma$  of  $H(x, x)$  is determined by the formula

$$\sigma = r - 2V(1, D_1, D_2, \dots, D_r). \quad (138)$$

All the remarks about the special cases that may occur, made for quadratic forms (§ 3), automatically carry over to hermitian forms.

DEFINITION 5: A hermitian form  $H(x, x) = \sum_{i,k=1}^n h_{ik} x_i \bar{x}_k$  is called positive (negative) semidefinite if for arbitrary values of the variables  $x_1, x_2, x_3, \dots, x_n$ , not all equal to zero,

$$H(x, x) \geq 0 \quad (\leq 0).$$

DEFINITION 6: A hermitian form  $H(x, x) = \sum_{i,k=1}^n h_{ik} x_i \bar{x}_k$  is called positive (negative) definite if for arbitrary values of the variables  $x_1, x_2, \dots, x_n$ , not all equal to zero,

$$H(x, x) > 0 \quad (< 0).$$

THEOREM 19: A hermitian form  $H(x, x) = \sum_{i,k=1}^n h_{ik} x_i \bar{x}_k$  is positive definite if and only if the following inequalities hold:

$$D_k = H \begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} > 0 \quad (k=1, 2, \dots, n). \quad (139)$$

THEOREM 20: A hermitian form  $H(x, x) = \sum_{i,k=1}^n h_{ik} x_i \bar{x}_k$  is positive semi-definite if and only if all the principal minors of  $H = \|h_{ik}\|_1^n$  are non-negative:

$$H \begin{pmatrix} i_1 & i_2 & \dots & i_p \\ i_1 & i_2 & \dots & i_p \end{pmatrix} \geq 0 \quad (140)$$

$$(i_1, i_2, \dots, i_p = 1, 2, \dots, n; p=1, 2, \dots, n).$$

The proofs of Theorems 19 and 20 are completely analogous to the proofs of Theorems 3 and 4 for quadratic forms.

The conditions for a hermitian form  $H(x, x)$  to be negative definite or semidefinite are obtained by applying (139) and (140) to the form  $-H(x, x)$ .

From Theorem 5' of Chapter IX (p. 274), we obtain the *Theorem on the reduction of a hermitian form to principal axes*:

THEOREM 21: Every hermitian form  $H(x, x) = \sum_{i,k=1}^n h_{ik} x_i \bar{x}_k$  can be reduced by a unitary transformation of the variables

$$x = U\xi \quad (UU^* = E) \quad (141)$$

to the canonical form

$$A(\xi, \xi) = \sum_{i=1}^n \lambda_i \xi_i \bar{\xi}_i, \quad (142)$$

where  $\lambda_1, \lambda_2, \dots, \lambda_n$  are the characteristic values of the matrix  $H = \|h_{ik}\|_1^n$ .

Theorem 21 follows from the formula

$$H = U \| \lambda_i \delta_{ik} \| U^{-1} = T^T \| \lambda_i \delta_{ik} \| \bar{T} \quad (U^T = \bar{U}^{-1} = T). \quad (143)$$

Let  $H(x, x) = \sum_{i,k=1}^n h_{ik} x_i \bar{x}_k$  and  $G(x, x) = \sum_{i,k=1}^n g_{ik} x_i \bar{x}_k$  be two hermitian forms. We shall study the pencil of hermitian forms  $H(x, x) - \lambda G(x, x)$

( $\lambda$  is a real parameter). This pencil is called *regular* if  $G(x, x)$  is positive definite. By means of the hermitian matrices  $H = \| h_{ik} \|_1^n$  and  $G = \| g_{ik} \|_1^n$  we form the equation

$$| H - \lambda G | = 0.$$

This equation is called the *characteristic equation of the pencil of hermitian forms*. Its roots are called the *characteristic values of the pencil*.

If  $\lambda_0$  is a characteristic value of the pencil, then there exists a column  $z = (z_1, z_2, \dots, z_n) \neq 0$  such that

$$Hz = \lambda_0 z.$$

We shall call the column  $z$  a *principal column* or *principal vector of the pencil*  $H(x, x) - \lambda G(x, x)$  corresponding to the characteristic value  $\lambda_0$ .

Then the following theorem holds:

**THEOREM 22:** *The characteristic equation of a regular pencil of hermitian forms  $H(x, x) - \lambda G(x, x)$  has  $n$  real roots  $\lambda_1, \lambda_2, \dots, \lambda_n$ . To these roots there correspond  $n$  principal vectors  $z^1, z^2, \dots, z^n$  satisfying the conditions of 'orthonormality':*

$$G(z^i, z^k) = \delta_{ik} \quad (i, k = 1, 2, \dots, n).$$

The proof is completely analogous to the proof of Theorem 8.

All extremal properties of the characteristic values of a regular pencil of quadratic forms remain valid for hermitian forms.

Theorems 10-17 remain valid if the term 'quadratic form' is replaced throughout by the term 'hermitian form.' The proofs of the theorems are then unchanged.

## § 10. Hankel Forms

**1.** Let  $s_0, s_1, \dots, s_{2n-2}$  be a sequence of numbers. We form, by means of these numbers, a quadratic form in  $n$  variables

$$S(x, y) = \sum_{i,k=0}^{n-1} s_{i+k} x_i x_k. \quad (144)$$

This is called a *Hankel form*. The matrix  $S = \| s_{i+k} \|_0^{n-1}$  corresponding to this form is called a *Hankel matrix*. It has the form

$$S = \begin{vmatrix} s_0 & s_1 & s_2 & \dots & s_{n-1} \\ s_1 & s_2 & s_3 & \dots & s_n \\ s_2 & s_3 & s_4 & \dots & s_{n+1} \\ \dots & \dots & \dots & \dots & \dots \\ s_{n-1} & s_n & s_{n+1} & \dots & s_{2n-2} \end{vmatrix}.$$

We denote the sequence of principal minors of  $S$  by  $D_1, D_2, \dots, D_n$ :

$$D_p = |s_{i+k}|_0^{p-1} \quad (p = 1, 2, \dots, n).$$

In this section we shall derive the fundamental results of Frobenius about the rank and signature of real Hankel forms.<sup>49</sup>

We begin by proving two lemmas.

**LEMMA 1:** *If the first  $h$  rows of the Hankel matrix  $S = \|s_{i+k}\|_0^{n-1}$  are linearly independent, but the first  $h+1$  rows linearly dependent, then*

$$D_h \neq 0.$$

*Proof.* We denote the first  $h+1$  rows of  $S$  by  $R_1, R_2, \dots, R_h, R_{h+1}$ . By assumption,  $R_1, R_2, \dots, R_h$  are linearly independent and  $R_{h+1}$  is expressed linearly in terms of them:

$$R_{h+1} = \sum_{j=1}^h \alpha_j R_{h-j+1}$$

or

$$s_q = \sum_{j=1}^h \alpha_j s_{q-j} \quad (q = h, h+1, \dots, h+n-1). \quad (145)$$

We write down the matrix formed from the first  $h$  rows  $R_1, R_2, \dots, R_h$  of  $S$ :

$$\begin{vmatrix} s_0 & s_1 & s_2 & \dots & s_{n-1} \\ s_1 & s_2 & s_3 & \dots & s_n \\ \dots & \dots & \dots & \dots & \dots \\ s_{h-1} & s_h & s_{h+1} & \dots & s_{h+n-2} \end{vmatrix}. \quad (146)$$

This matrix is of rank  $h$ . On the other hand, by (145) every column of the matrix can be expressed linearly in terms of the preceding  $h$  columns and hence in the terms of the first  $h$  columns. But since the rank of (146) is  $h$ , these first  $h$  columns of (146) must then be linearly independent, i.e.,

$$D_h \neq 0.$$

This proves the lemma.

<sup>49</sup> See [162].

LEMMA 2: If in the matrix  $S = \| s_{i+k} \|_0^{n-1}$ , for a certain  $h$  ( $< n$ ),

$$D_h \neq 0, \quad D_{h+1} = \dots = D_n = 0 \quad (147)$$

and

$$t_{ik} = \frac{S \begin{pmatrix} 1 & \dots & h & h+i+1 \\ 1 & \dots & h & h+k+1 \end{pmatrix}}{S \begin{pmatrix} 1 & \dots & h \\ 1 & \dots & h \end{pmatrix}} = \frac{1}{D_h} \begin{vmatrix} D_h & s_{h+k} \\ & \vdots \\ & s_{2h+k-1} \\ s_{h+i} & \dots & s_{2h+i-1} & s_{2h+i+k} \end{vmatrix}, \quad (148)$$

$$(i, k, = 0, 1, \dots, n-h-1)$$

then the matrix  $T = \| t_{ik} \|_0^{n-h-1}$  is also a Hankel matrix and all its elements above the second diagonal are zero, i.e., there exist numbers  $t_{n-h-1}, \dots, t_{2n-2h-2}$  such that

$$t_{ik} = t_{i+k} \quad (i, k = 0, 1, \dots, n-h-1; t_0 = t_1 = \dots = t_{n-h-2} = 0).$$

*Proof.* We introduce the matrices

$$T_p = \| t_{ik} \|_0^{p-1} \quad (p = 1, 2, \dots, n-h).$$

In this notation  $T = T_{n-h}$ .

We shall show that every  $T_p$  ( $p = 1, 2, \dots, n-h$ ) is a Hankel matrix and that  $t_{ik} = 0$  for  $i+k \leq p-2$ . The proof is by induction with respect to  $p$ .

For the matrix  $T_1$ , our assertion is trivial; for  $T_2$ , it is obvious, since

$$T_2 = \begin{vmatrix} t_{00} & t_{01} \\ t_{10} & t_{11} \end{vmatrix}, \quad t_{01} = t_{10} \quad (\text{because } S \text{ is symmetric}) \text{ and } t_{00} = \frac{D_{h+1}}{D_h} = 0.$$

Let us assume that our assertion is true for the matrices  $T_p$  ( $p < n-h$ ) we shall show that it is also true for  $T_{p+1} = \| t_{ik} \|_0^p$ . From the assumption it follows that there exist numbers  $t_{p-1}, t_p, \dots, t_{2p-2}$  such that with  $t_0 = \dots = t_{p-2} = 0$

$$T_p = \| t_{i+k} \|_0^{p-1}.$$

Here

$$|T_p| = t_{p-1}^p. \quad (149)$$

On the other hand, using Sylvester's determinant identity (see (28) or page 32), we find:

$$|T_p| = \frac{D_{h+p}}{D_h} = 0. \quad (150)$$

Comparing (149) with (150), we obtain

$$t_{p-1} = 0. \quad (151)$$

Furthermore from (148)

$$t_{ik} = s_{2h+i+k} + \frac{1}{D_h} \begin{vmatrix} & & s_{h+k} \\ & & \cdot \\ & & \cdot \\ D_h & & \cdot \\ & & s_{2h+k-1} \\ s_{h+i} \cdots s_{2h+i-1} & 0 \end{vmatrix}. \quad (152)$$

By the preceding lemma, it follows from (147) that the  $(h+1)$ -th row of the matrix  $S = \|s_{i+k}\|_0^{n-1}$  is linearly dependent on the first  $h$  rows:

$$s_q = \sum_{g=1}^h \alpha_g s_{q-g} \quad (q = h, h+1, \dots, h+n-1). \quad (153)$$

Let  $i, k \leq p \leq i+k \leq 2p-1$ . Then one of the numbers  $i$  or  $k$  is less than  $p$ . Without loss of generality, we assume that  $i < p$ . Then, when we expand, by (153), the last column of the determinant of the right-hand side of (152) and use the relations (152) again, we shall have

$$\begin{aligned} t_{ik} &= s_{2h+i+k} + \sum_{g=1}^h \frac{\alpha_g}{D_h} \begin{vmatrix} & & s_{h+k-g} \\ & & \cdot \\ & & \cdot \\ D_h & & \cdot \\ & & s_{2h+k-g-1} \\ s_{h+i} \cdots s_{2h+i-1} & 0 \end{vmatrix} \\ &= s_{2h+i+k} + \sum_{g=1}^h \alpha_g (t_{i, k-g} - s_{2h+i+k-g}). \end{aligned} \quad (154)$$

By the induction hypothesis (151) holds, and since in (154)  $i < p, k-g < p$  and  $i+k-g \leq 2p-2$ , we have  $t_{i, k-g} = t_{i+k-g}$ . Therefore, for  $i+k < p$  all the  $t_{ik} = 0$ , and for  $p \leq i+k \leq 2p-1$  the value of  $t_{ik}$ , by (154), depends on  $i+k$  only.

Thus,  $T_{p+1}$  is a Hankel matrix, and all its elements  $t_0, t_1, \dots, t_{p-1}$  above the second diagonal are zero.

This proves the lemma.

Using Lemma 2, we shall prove the following theorem:



THEOREM 23: If the Hankel matrix  $S = \|s_{i+k}\|_0^{n-1}$  has rank  $r$  and if for some  $h$  ( $< r$ )

$$D_h \neq 0, \quad D_{h+1} = \dots = D_r = 0,$$

then the principal minor of order  $r$  formed from the first  $h$  and the last  $r - h$  rows and columns of  $S$  is not zero:

$$D^{(r)} = S \begin{pmatrix} 1 & \dots & h & n-r+h+1 & n-r+h+2 & \dots & n \\ 1 & \dots & h & n-r+h+1 & n-r+h+2 & \dots & n \end{pmatrix} \neq 0.$$

*Proof.* By the preceding lemma, the matrix

$$T = \|t_{ik}\|_0^{n-h-1} \quad \left( t_{ik} = \frac{S \begin{pmatrix} 1 & \dots & h & h+i+1 \\ 1 & \dots & h & h+k+1 \end{pmatrix}}{S \begin{pmatrix} 1 & \dots & h \\ 1 & \dots & h \end{pmatrix}} \quad (i, k = 0, 1, \dots, n-h+1) \right)$$

is a Hankel matrix in which all the elements above the second diagonal are zero. Therefore

$$|T| = t_{0, n-h-1}^{n-h}.$$

On the other hand,<sup>50</sup>  $|T| = \frac{D_n}{D_h} = 0$ . Therefore  $t_{0, n-h-1} = 0$ , and the matrix  $T$  has the form

$$T = \begin{vmatrix} 0 & \dots & \dots & \dots & \dots & 0 \\ & \ddots & & & & u_{n-h-1} \\ & & \ddots & & & \vdots \\ & & & \ddots & & \vdots \\ & & & & \ddots & \vdots \\ & & & & & u_2 \\ & & & & & \vdots \\ 0 & u_{n-h-1} & \dots & u_2 & u_1 \end{vmatrix}$$

The rank of  $T$  must be  $r - h$ .<sup>51</sup> Therefore for  $r < n - 1$  in the matrix  $T$  the elements  $u_{r-h+1} = \dots = u_{n-h+1} = 0$ , and

<sup>50</sup> By Sylvester's determinant identity (see (28) on p. 32).

<sup>51</sup> From Sylvester's identity it follows that all the minors of  $T$  in which the order exceeds  $r - h$  are zero. On the other hand,  $S$  contains some non-vanishing minors of order  $r$  bordering  $D_h$ . Hence it follows that the corresponding minor of order  $r - h$  of  $T$  is different from zero.

$$T = \begin{vmatrix} 0 & \dots & \dots & \dots & \dots & 0 \\ \vdots & & & & & \vdots \\ \vdots & & & & & \vdots \\ \vdots & & & & & \vdots \\ \vdots & & & & & \vdots \\ 0 & \dots & 0 & u_{r-h} & \dots & u_1 \end{vmatrix} \quad (u_{r-h} \neq 0).$$

But then, by Sylvester's identity (see page 32),

$$D^{(r)} = D_h T \begin{pmatrix} n-r+1 & \dots & n-h \\ n-r+1 & \dots & n-h \end{pmatrix} = D_h u_{r-h}^{r-h} \neq 0,$$

and this is what we had to prove.

Let us consider a real<sup>52</sup> Hankel form  $S(x, x) = \sum_{i,k=0}^{n-1} s_{i+k} x_i x_k$  of rank  $r$ . We denote by  $\pi$ ,  $\nu$ , and  $\sigma$ , respectively, the number of positive and of negative squares and the signature of the form:

$$\pi + \nu = r, \quad \sigma = \pi - \nu = r - 2\nu.$$

By the theorem of Jacobi (p. 303) these values can be determined from the signs of the successive minors

$$D_0 = 1, D_1, D_2, \dots, D_{r-1}, D_r \quad (155)$$

by the formulas

$$\left. \begin{aligned} \pi &= P(1, D_1, \dots, D_r), & \nu &= V(1, D_1, \dots, D_r), \\ \sigma &= P(1, D_1, \dots, D_r) - V(1, D_1, \dots, D_r) = r - 2V(1, D_1, \dots, D_r). \end{aligned} \right\} \quad (156)$$

These formulas become inapplicable when the last term in (155) or any three consecutive terms are zero (see § 3). However, as Frobenius has shown, for Hankel forms there is a rule that enables us to use the formulas (156) in the general case:

**THEOREM 24 (Frobenius):** *For a real Hankel form  $S(x, x) = \sum_{i,k=0}^{n-1} s_{i+k} x_i x_k$  of rank  $r$  the values of  $\pi$ ,  $\nu$ , and  $\sigma$  can be determined by the formulas (156) provided that*

<sup>52</sup> In the preceding Lemmas 1 and 2 and in Theorem 23, the ground field can be taken as an arbitrary number field—in particular, the field of complex or of real numbers.

1) for

$$D_h \neq 0, D_{h+1} = \dots = D_r = 0 \quad (h < r) \quad (157)$$

 $D_r$  is replaced by  $D^{(r)}$ , where

$$D^{(r)} = S \begin{pmatrix} 1 & \dots & h & n-r+h+1 & \dots & n \\ 1 & \dots & h & n-r+h+1 & \dots & n \end{pmatrix} \neq 0;$$

2) in any group of  $p$  consecutive zero determinants

$$(D_h \neq 0) \quad D_{h+1} = D_{h+2} = \dots = D_{h+p} = 0 \quad (D_{h+p+1} \neq 0) \quad (158)$$

a sign is attributed to the zero determinants according to the formula

$$\text{sign } D_{h+j} = (-1)^{\frac{j(j-1)}{2}} \text{sign } D_h. \quad (159)$$

The values of  $P$ ,  $V$ , and  $P - V$  corresponding to the group (158) are then:<sup>53</sup>

	$p$ odd	$p$ even	
$P_{h,p} = P(D_h, D_{h+1}, \dots, D_{h+p+1})$	$\frac{p+1}{2}$	$\frac{p+1+\varepsilon}{2}$	
$V_{h,p} = V(D_h, D_{h+1}, \dots, D_{h+p+1})$	$\frac{p+1}{2}$	$\frac{p+1-\varepsilon}{2}$	
$P_{h,p} - V_{h,p}$	0	$\varepsilon$	

$$\varepsilon = (-1)^{\frac{p}{2}} \text{sign } \frac{D_{h+p+1}}{D_h}. \quad (160)$$

*Proof.* To begin with we consider the case where  $D_r \neq 0$ . Then the forms  $S(x, x) = \sum_{i,k=0}^{n-1} s_{i+k} x_i x_k$  and  $S_r(x, x) = \sum_{i,k=0}^{r-1} s_{i+k} x_i x_k$  have not only the same rank  $r$ , but also the same signature  $\sigma$ . For let  $S(x, x) = \sum_{i=1}^r \varepsilon_i Z_i^2$ , where the  $Z_i$  are real linear forms and  $\varepsilon_i = \pm 1$  ( $i = 1, 2, \dots, r$ ). We set  $x_{r+1} = \dots = x_{n-1} = 0$ . Then the forms  $S(x, x)$  and  $Z_i$  go over, respectively, into  $S_r(x, x)$  and  $\widehat{Z}_i$  ( $i = 1, 2, \dots, r$ ); and  $S_r(x, x) = \sum_{i=1}^r \varepsilon_i \widehat{Z}_i^2$ , i.e.,  $S_r(x, x)$  has

<sup>53</sup> The formulas (159) and (160) are also applicable to (157), but we have to set  $p = r - h - 1$  and interpret  $D_{h+p+1}$  not as  $D_r = 0$ , but as  $D^{(r)} \neq 0$ .

the same number of positive and negative squares as  $S(x, x)$ .<sup>54</sup> Thus the signature of  $S_r(x, x)$  is  $\sigma$ .

We now vary the parameters  $s_0, s_1, \dots, s_{2r-2}$  continuously in such a way that for the new parameter values  $s_0^*, s_1^*, \dots, s_{2r-2}^*$  all the terms of the sequence<sup>55</sup>

$$1, D_1^*, D_2^*, \dots, D_r^* \quad (D_q^* = |s_{i+k}^*|_{q-1}^{q-1}; q = 1, 2, \dots, r)$$

are different from zero and that in the process of variation none of the non-zero determinants (155) vanishes.<sup>56</sup>

Since the rank of  $S_r(x, x)$  does not change during the variation, its signature also remains unchanged (see p. 309). Therefore

$$\sigma = P(1, D_1^*, \dots, D_r^*) - V(1, D_1^*, \dots, D_r^*). \quad (161)$$

If  $D_i \neq 0$  for some  $i$ , then  $\text{sign } D_i^* = \text{sign } D_i$ . Therefore the whole problem reduces to determining the variations in sign among those  $D_i^*$  that correspond to  $D_i = 0$ . More accurately, for every group of the form (158) we have to determine

$$P(D_h^*, D_{h+1}^*, \dots, D_{h+p+1}^*) - V(D_h^*, D_{h+1}^*, \dots, D_{h+p}^*, D_{h+p+1}^*).$$

For this purpose we set:

$$t_{ik} = \frac{1}{D_h} \begin{vmatrix} & & & s_{h+k} \\ & & & \cdot \\ & & & \cdot \\ & & & \cdot \\ & & s_{2h+k-1} & \\ s_{h+i} \cdots s_{2h+i-1} & s_{2h+i+k} \end{vmatrix} \quad (i, k = 0, 1, \dots, p).$$

By Lemma 2, the matrix  $T = \|t_{ik}\|_0^p$  is a Hankel matrix and all its elements above the second diagonal are zero, so that  $T$  has the form

<sup>54</sup> The linear forms  $\widehat{Z}_1, \widehat{Z}_2, \dots, \widehat{Z}_r$  are linearly independent, because the quadratic form  $S(x, x) = \sum_{i=1}^r \varepsilon_i \widehat{Z}_i^2$  is of rank  $r$  ( $D_r \neq 0$ ).

<sup>55</sup> In this section, the asterisk  $*$  does not indicate the adjoint matrix.

<sup>56</sup> Such a variation of the parameter is always possible, because in the space of the parameters  $s_0, s_1, \dots, s_{2r-2}$  an equation of the form  $D_i = 0$  determines a certain algebraic hypersurface. If a point lies in some such hypersurfaces, then it can always be approximated by arbitrarily close points that do not lie in these hypersurfaces.

$$T = \begin{vmatrix} 0 & \dots & 0 & t_p \\ \vdots & & \vdots & * \\ \vdots & & \vdots & \vdots \\ 0 & \vdots & & \vdots \\ t_p & * & \dots & * \end{vmatrix}. \quad (162)$$

We denote the successive minors of  $T$  by  $\widehat{D}_1, \widehat{D}_2, \dots, \widehat{D}_{p+1}$ :

$$\widehat{D}_q = |t_{ik}|_0^{q-1} \quad (q = 1, 2, \dots, p+1).$$

Together with  $T$ , we consider the matrix

$$T^* = \|t_{ik}^*\|_0^p,$$

where

$$t_{ik}^* = \frac{1}{D_h^*} \begin{vmatrix} & & s_{h+k}^* \\ & D_h^* & \vdots \\ & & s_{2h+k-1}^* \\ s_{h+i}^* \dots s_{2h+i-1}^* & & s_{2h+i+k}^* \end{vmatrix} \quad (i, k = 0, 1, \dots, p)$$

and the corresponding determinants

$$\widehat{D}_q^* = |t_{ik}^*|_0^{q-1} \quad (q = 1, 2, \dots, p+1).$$

By Sylvester's determinant identity,

$$D_{h+q}^* = D_h^* \widehat{D}_q^* \quad (q = 1, 2, \dots, p+1).$$

Therefore

$$\begin{aligned} P(D_h^*, D_{h+1}^*, \dots, D_{h+p+1}^*) - V(D_h^*, D_{h+1}^*, \dots, D_{h+p+1}^*) \\ = P(1, \widehat{D}_1^*, \dots, \widehat{D}_{p+1}^*) - V(1, \widehat{D}_1^*, \dots, \widehat{D}_{p+1}^*) = \widehat{\sigma}^*, \end{aligned} \quad (163)$$

where  $\widehat{\sigma}^*$  is the signature of the form

$$T^*(x, x) = \sum_{i, k=0}^p t_{ik}^* x_i x_k.$$

Together with  $T^*(x, x)$ , we consider the forms

$$T(x, x) = \sum_{i, k=0}^p t_{i+k} x_i x_k \quad \text{and} \quad T^{**}(x, x) = t_p(x_0 x_p + x_1 x_{p-1} + \dots + x_p x_0).$$

The matrix  $T^{**}$  is obtained from  $T$  (see (162)) when we replace in the latter all the elements above the second diagonal by zeros. We denote the signatures of  $T(x, x)$  and  $T^{**}(x, x)$  by  $\hat{\sigma}$  and  $\hat{\sigma}^{**}$ . Since  $T^*(x, x)$  and  $T^{**}(x, x)$  are obtained from  $T(x, x)$  by variations of the coefficients during which the rank of the form does not change ( $|T^{**}| = |T| = \frac{D_{h+p+1}}{D_h} \neq 0$ ,  $|T^*| = \frac{D_{h+p+1}}{D_h} \neq 0$ ), the signatures of  $T(x, x)$ ,  $T^*(x, x)$ , and  $T^{**}(x, x)$  must also be equal:

$$\hat{\sigma} = \hat{\sigma}^* = \hat{\sigma}^{**}. \quad (164)$$

But

$$T^{**}(x, x) = \begin{cases} 2t_p(x_0x_{2k-1} + \dots + x_{k-1}x_k) & \text{for odd } p, \\ t_p[2(x_0x_{2k} + \dots + x_{k-1}x_{k+1}) + x_k^2] & \text{for even } p \end{cases}$$

Since every product of the form  $x_\alpha x_\beta$  with  $\alpha \neq \beta$  can be replaced by a difference of squares  $\left(\frac{x_\alpha + x_\beta}{2}\right)^2 - \left(\frac{x_\alpha - x_\beta}{2}\right)^2$ , we can obtain a decomposition of  $T^{**}(x, x)$  into independent real squares and we have

$$\hat{\sigma}^{**} = \begin{cases} 0 & \text{for odd } p, \\ \text{sign } t_p & \text{for even } p. \end{cases} \quad (165)$$

On the other hand, from (162),

$$\frac{D_{h+p+1}}{D_h} = |T| = (-1)^{\frac{p(p+1)}{2}} t_p^{p+1}. \quad (166)$$

From (163), (164), (165), and (166), it follows that:

$$\begin{aligned} P(D_h^*, D_{h+1}^*, \dots, D_{h+p+1}^*) - V(D_h^*, D_{h+1}^*, \dots, D_{h+p+1}^*) \\ = \begin{cases} 0 & \text{for odd } p, \\ \varepsilon & \text{for even } p. \end{cases} \end{aligned} \quad (167)$$

where

$$\varepsilon = (-1)^{\frac{p}{2}} \text{sign } \frac{D_{h+p+1}}{D_h}$$

Since

$$P(D_{h+1}^*, D_{h+2}^*, \dots, D_{h+p+1}^*) + V(D_{h+1}^*, D_{h+2}^*, \dots, D_{h+p+1}^*) = p+1, \quad (168)$$

the table (160) can be deduced from (167) and (168).

Now let  $D_r = 0$ . Then for some  $h < r$

$$D_h \neq 0, D_{h+1} = \dots = D_r = 0.$$

In this case, by Theorem 25,

$$D^{(r)} = S \begin{pmatrix} 1 \dots h & n-r+h+1 & \dots n \\ 1 \dots h & n-r+h+1 & \dots n \end{pmatrix} \neq 0.$$

The case to be considered reduces to the preceding case by renumbering the variables in the quadratic form  $S(x, x) = \sum_{i, k=0}^{n-1} s_{i+k} x_i x_k$ . We set:

$$\begin{aligned} \tilde{x}_0 = x_0, \quad \dots, \quad \tilde{x}_{h-1} = x_{h-1}, \quad \tilde{x}_h = x_{n-r+h}, \quad \dots, \quad \tilde{x}_{r-1} = x_{n-1}, \\ \tilde{x}_r = x_h, \quad \dots, \quad \tilde{x}_{n-1} = x_{n-r+h-1}. \end{aligned}$$

$$\text{Then } S(x, x) = \sum_{i, k=0}^{n-1} \tilde{s}_{i+k} x_i x_k.$$

Starting from the structure of the matrix  $T$  on page 346 and using the relations

$$\hat{D}_j = \frac{D_{h+j}}{D_h}, \quad \hat{\tilde{D}}_j = \frac{\tilde{D}_{h+j}}{\tilde{D}_h} \quad (j=1, 2, \dots, n-h)$$

obtained from Sylvester's determinant identity, we find that the sequence  $1, \tilde{D}_1, \tilde{D}_2, \dots, \tilde{D}_n$  is obtained from  $1, D_1, D_2, \dots, D_n$  by replacing the single element  $D_r$  by  $D^{(r)}$ .

We leave it to the reader to verify that the table (160) corresponds to the attribution of signs to the zero determinants given by (159).

This completes the proof of the theorem.

*Note.* It follows from (166) that for odd  $p$  ( $p$  is the number of zero determinants in the group (158))

$$\text{sign } \frac{D_{h+p+1}}{D_h} = (-1)^{\frac{p+1}{2}}$$

In particular, for  $p=1$  we have  $D_h D_{h+2} < 0$ . In this case, we can omit  $D_{h+1}$  in computing  $V(1, D_1, \dots, D_r)$ , thus obtaining Gundenfinger's rule. In exactly the same way, we obtain Frobenius' rule (see page 304) from (160) for  $p=2$ .

# INDEX

[Numbers in *italics* refer to Volume Two]

- ABSOLUTE CONCEPTS, 184
- Addition of congruences, 182
- Addition of operators, 57
- Adjoint matrix, 82
- Adjoint operator, 265
- Algebra, 17
- Algorithm of Gauss, 23ff.  
    generalized, 45
- Angle between vectors, 242
- Axes, principal, 309  
    reduction to, 309
- BASIS (ES), 51  
    characteristic, 73  
    coordinates of vector in, 53  
    Jordan, 201  
        lower, 202  
    orthonormal, 242, 245
- Bessel, inequality of, 259
- Bézout, generalized theorem of, 81
- Binet-Cauchy formula, 9
- Birkhoff, G. D., 147
- Block, of matrix, 41  
    diagonal, isolated, 75  
    Jordan, 151
- Block multiplication of matrices, 42
- Bundle of vectors, 183
- Bunyakovskii's inequality, 255
- CARTAN, theorem of, 4
- Cauchy, formula of Binet-, 9  
    system of, 115
- Cauchy identity, 10
- Cauchy index, 174, 216
- Cayley, formulas of, 279
- Cayley-Hamilton theorem, 83, 197.
- Cell, of matrix, 41
- Chain, *see* Jordan, Markov, Sturm
- Characteristic basis, 73
- Characteristic direction, 71
- Characteristic equation, 70, 310, 338
- Characteristic matrix, 82
- Characteristic polynomial, 71, 82
- Characterization of root, minimal, 319  
    maximal-minimal, 321, 322
- Chebyshev, 173, 240  
    polynomials of, 259
- Chebyshev-Markov, formula of, 248  
    theorem of, 247
- Chetaev, 121
- Chipart, 173, 221
- Coefficients of Fourier, 261
- Coefficients of influence, reduced, 111
- Column, principal, 338
- Column matrix, 2
- Columns, Jordan chains of, 165
- Components, of matrix, 105  
    of operator, hermitian, 268  
    skew-symmetric, 281  
    symmetric, 281
- Compound matrix, 19ff., 20
- Computation of powers of matrix, 109
- Congruences, 181, 182
- Constraint, 320
- Convergence, 110, 112
- Coordinates, transformation of, 59  
    of vector, 53
- Coordinate transformation, matrix of, 60'
- D'ALEMBERT-EULER, theorem of, 286
- Danilevskii, 214
- Decomposition, of matrix into triangular  
    factors, 33ff.  
    polar, of operator, 276, 286; 6  
    of space, 248
- Defect of vector space, 64
- Derivative, multiplicative, 133
- Determinant identity of Sylvester, 32, 33
- Determinant of square matrix, 1
- Diagonal matrix, 3
- Dilatation of space, 287
- Dimension, of matrix, 1  
    of vector space, 51
- Direction, characteristic, 71
- Discriminant of form, 333



- Divisors, elementary, 142, 144, 194  
 admissible, 238  
 geometrical theory of, 175  
 infinite, 27
- Dmitriev, 87
- Domain of stability, 232
- Dynkin, 87
- EIGENVALUE, 69
- Elements of matrix, 1
- Elimination method of Gauss, 23ff.
- Equivalence, of matrices, 61, 132, 133  
 of pencils, strict, 24
- Ergodic theorem for Markov chains, 95
- Erugin, theorem of, 122
- Euler-D'Alembert, theorem of, 286
- FACTOR SPACE, 183
- Faddeev, method of, 87
- Field, 1
- Forces, linear superposition of, 28
- Form, bilinear, 294  
 Hankel, 338; 205  
 hermitian, 244, 331  
 bilinear, 332  
 canonical form of, 337  
 negative definite, 337  
 negative semidefinite, 336  
 pencil of, *see* pencil  
 positive definite, 337  
 positive semidefinite, 336  
 rank of, 333  
 signature of, 334  
 singular, 333  
 quadratic, 246, 294  
 definite, 305  
 discriminant of, 294  
 rank of, 296  
 real, 294  
 reduction of, 299ff.  
 reduction to principal axes, 309  
 restricted, 306  
 semidefinite, 304  
 signature of, 296, 298  
 singular, 294
- Fourier series, 261
- Frobenius, 304, 339, 343; 53  
 theorem of, 343; 53
- Function, entire, 169  
 left value of, 81
- GANTMACHER, 103
- Gauss, algorithm of, 23ff.  
 generalized, 45  
 elimination method of, 23ff.
- Gaussian form of matrix, 39
- Golubchikov, 124
- Governors, 172, 233
- Gram, criterion of, 247
- Gramian, 247, 251
- Group, 18  
 unitary, 268
- Gundenfinger, 304
- HADAMARD INEQUALITY, 252  
 generalized, 254
- Hamilton-Cayley theorem, 83, 197
- Hankel form, 338; 205
- Hankel matrix, 338; 205
- Hermite, 172, 202, 210
- Hermite-Biehler theorem, 228
- Hurwitz, 173, 190, 210
- Hurwitz matrix, 190
- Hyperlogarithm, 169
- IDENTITY OPERATOR, 66
- Imprimitivity, index of, 80
- Ince, 147
- Inertia, law of, 297, 334
- Integral, multiplicative, 132, 138  
 product, 132
- Invariant plane, of operator, 283
- JACOBI, formula of, 302, 336  
 identity of, 114  
 method of, 300  
 theorem of, 303
- Jacobi matrix, 99
- Jordan basis, 201
- Jordan block, 151
- Jordan chains of columns, 165
- Jordan form of matrix, 152, 201, 202
- Jordan matrix, 152, 201
- KARPELEVICH, 87
- Kernel of  $\lambda$ -matrix, 39
- Kolmogorov, 83, 87, 92
- Kotelyanskii, 103  
 lemma of, 71
- Krein, 221, 250
- Kronecker, 75; 25, 37, 40
- Krylov, 203  
 transformation of, 206
- LAGRANGE, method of, 299
- Lagrange interpolation polynomial, 101
- Lagrange-Sylvester interpolation polynomial, 97
- $\lambda$ -matrix, 130  
 kernel of, 39

- Lappo-Danilevskii, 168, 170, 171  
 Left value, 81  
 Legendre polynomials, 258  
 Liénard, 173, 221  
 Liénard-Chipart stability criterion, 221  
 Limit of sequence of matrices, 33  
 Linear (in)dependence of vectors, 51  
 Linear transformation, 3  
 Logarithm of matrix, 239  
 Lyapunov, 173, 185  
     criterion of, 180  
     equivalence in the sense of, 118  
     theorem of, 187  
 Lyapunov matrix, 117  
 Lyapunov transformation, 117  
 MACMILLAN, 115  
 Mapping, affine, 245  
 Markov, 173, 240  
     theorem of, 242  
 Markov chain, acyclic, 88  
     cyclic, 88  
     fully regular, 88  
     homogeneous, 88  
         period of, 96  
     (ir)reducible, 88  
     regular, 88  
 Markov parameters, 233, 234  
 Matricant, 127  
 Matrices, addition of, 4  
     group property, 18  
     annihilating polynomial of, 89  
     applications to differential equations, 116ff.  
     congruence of, 296  
     difference of, 5  
     equivalence of, 132, 133  
     equivalent, 61ff.  
     left-equivalence of, 132, 133  
     limit of sequence of, 33  
     multiplication on left by  $H$ , 14  
     product of, 6  
     quotient of, 17  
     rank of product, 12  
     similarity of, 67  
     unitary similarity of, 242  
     with same real part of spectrum, 122  
     adjoint, 82, 266  
         reduced, 90  
     blocks of, 41  
     canonical form of, 63, 135, 136, 139, 141, 152, 192, 201, 202, 264, 265  
     cells of, 41  
     characteristic, 82  
     characteristic polynomial of, 82  
     Matrix, column, 2  
         commuting, 7  
         companion, 149  
         completely reducible, 81  
         complex, 1ff.  
             orthogonal, normal form of, 23  
             representation of as product, 6  
             skew-symmetric, normal form of, 18  
             symmetric, normal form of, 11  
         components of, 105  
         compound, 19ff., 20  
         computation of power of, 109  
         constituent, 105  
         of coordinate transformation, 60  
         cyclic form of, 54  
         decomposition into triangular factors, 33ff.  
         derivative of, 117  
         determinant of, 1, 5  
         diagonal, 3  
             multiplication by, 8  
         diagonal form of, 152  
         dimension of, 1  
         elementary, 132  
         elementary divisors of, 142, 144, 194  
         elements of, 1  
         function of, 95ff.  
             defined on spectrum, 96  
         fundamental, 73  
         Gaussian form of, 39  
         Hankel, 338; 205  
             projective, 20  
         Hurwitz, 190  
         idempotent, 226  
         infinite, rank of, 239  
         integral, 126; 113  
             normalized, 114  
         invariant polynomials of, 139, 144, 194  
         inverse of, 15  
         minors of, 19ff.  
         irreducible, 50  
             (im)primitive, 80  
         Jacobi, 99  
         Jordan form of, 152, 201, 202  
          $\lambda$ , 130  
         and linear operator, 56  
         logarithm of, 239  
         Lyapunov, 117  
         minimal polynomial of, 89  
         uniqueness of, 90  
         minor of, 2  
         principal, 2  
         multiplication of, by number, 5  
         by matrix, 17

- Matrix**, nilpotent, 226  
     non-negative, 50  
         totally, 98  
     non-singular, 15  
     normal, 269  
     normal form of, 150, 192, 201, 202  
     notation for, 1  
     order of, 1  
     orthogonal, 263  
     oscillatory, 103  
     partitioned, 41, 42  
     permutable, 7  
     permutation of, 50  
     polynomial, *see* polynomial matrix  
     polynomials in, permutability of, 13  
     positive, 50  
         spectra of, 53  
         totally, 98  
     power of, 12  
         computation of, 109  
     power series in, 113  
     principal minor of, 2  
     quasi-triangular, 43  
     rank of, 2  
     reducible, 50, 51  
         normal form of, 75  
     representation as product, 264  
     root of non-singular, 233  
     root of singular, 234ff., 239  
     Routh, 191  
     row, 2  
     of simple structure, 73  
     singular, 15  
     skew-symmetric, 19  
     square, 1  
     square root of, 239  
     stochastic, 83  
         fully regular, 88  
         regular, 88  
     spur of, 87  
     subdiagonal of, 13  
     superdiagonal of, 13  
     symmetric, 19  
     trace of, 87  
     transformation of coordinate, 60  
     transforming, 35, 60  
     transpose of, 19  
     triangular, 18, 218; 155  
     unit, 12  
     unitary, 263, 269  
     unitary, representation of as product, 5  
     upper quasi-triangular, 43  
     upper triangular, 18  
**Matrix addition**, properties of, 4  
**Matrix equations**, 215ff.  
     uniqueness of solution, 16  
**Matrix multiplication**, 6, 7  
**Matrix polynomials**, 76  
     left quotient of, 78  
     multiplication of, 77  
**Maxwell**, 172  
**Mean**, convergence in, of series, 260  
**Metric**, 242  
     euclidean, 245  
     hermitian, 243, 244  
         positive definite, 243.  
         positive semidefinite, 243  
**Minimal indices for columns**, 38  
**Minor**, 2  
     almost principal, 102  
     of zero density, 104  
**Modulus**, left, 275  
**Moments**, problem of, 236, 237  
**Motion**, of mechanical system, 12  
     of point, 121  
     stability of, 125  
     asymptotic, 125  
**NAÏMARK**, 221, 233, 250  
**Nilpotency**, index of, 226  
**Norm**, left, 275  
     of vector, 243  
**Null vector**, 52  
**Nullity of vector space**, 64  
**Number space**,  $n$ -dimensional, 52  
**OPERATIONS**, elementary, 134  
**Operator (linear)**, 55, 66  
     adjoint, 265  
     decomposition of, 281  
     hermitian, 268  
         positive definite, 274  
         positive semidefinite, 274  
         projective, 20  
         spectrum of, 272  
     identity, 66  
     invariant plane of, 283  
     matrix corresponding to, 56  
     normal, 268  
     positive definite, 280  
     positive semidefinite, 280  
     normal, 280  
     orthogonal, of first kind, 281  
         (im)proper, 281  
         of second kind, 281  
     polar decomposition of, 276, 286  
     real, 282  
     semidefinite, 274, 280

- Operator (linear), of simple structure, 72
  - skew-symmetric, 280
  - square root of, 275
  - symmetric, 280
  - transposed, 280
  - unitary, 268
    - spectrum of, 273
- Operators, addition of, 57
  - multiplication of, 58
- Order of matrix, 1
- Orlando, formula of, 196
- Orthogonal complement, 266
- Orthogonalization, 256
- Oscillations, small,  $\pi$  system, 326
- PARAMETERS, homogeneous, 26
  - Markov, 233, 234
- Parseval, equality of, 261
- Peano, 127
- Pencil of hermitian forms, 338
  - characteristic equation of, 338
  - characteristic values of, 338
  - principal vector of, 338
- Pencil(s) of matrices, canonical form of, 37, 39
  - congruent, 41
  - elementary divisors of, infinite, 27
  - rank of, 29
  - regular, 25
  - singular, 25
  - strict equivalence of, 24
- Pencil of quadratic forms, 310
  - characteristic equation of, 310
  - characteristic value of, 310
  - principal column of, 310
  - principal matrix of, 312
  - principal vector of, 310
- Period, of Markov chain, 96
- Permutation of matrix, 50
- Perron, 53
  - formula of, 116
- Petrovskii, 113
- Polynomial(s), annihilating, 176, 177
  - minimal, 176
    - of square matrix, 89
  - of Chebyshev, 259
  - characteristic, 71
  - interpolation, 97, 101, 103
  - invariant, 139, 144, 194
  - of Legendre, 258
  - matrix, *see* matrix polynomials
  - minimal, 89, 176, 177
  - monic, 176
  - scalar, 76
  - positive pair of, 227
- Polynomial matrix, 76, 130
  - elementary operations on, 130, 131
  - regular, 76
  - order of, 76
- Power of mat.  $x$ , 12
- Probability, absolute, 93
  - limiting, 94
  - mean limiting, 96
- transition, 82
  - final, 88
  - limiting, 88
  - mean limiting, 96
- Product, inner, of vectors, 243
  - scalar, of vectors, 242, 243
  - of operators, 58
  - of sequences, 6
- Pythagoras, theorem of, 244
- QUASI-ERGODIC THEOREM, 95
- Quasi-triangular matrix, 43
- Quotients of matrices, 17
- RANK, of infinite matrix, 239
  - of matrix, 2
  - of pencil, 29
  - of vector space, 64
- Relative concepts, 184
- Right value, 81
- Ring, 17
- Romanovskii, 83
- Root of matrix, 233, 234ff., 239
- Rotation of space, 287
- Routh, 173, 201
  - criterion of, 180
- Routh-Hurwitz, criterion of, 194
- Routh matrix, 191
- Routh scheme, 179
- Row matrix, 2
- SCHLESINGER, 133
- Schur, formulas of, 46
- Schwarz, inequality of, 255
- Sequence of vectors, 256, 260
- Series, convergence of, 260
  - fundamental, of solutions, 38
- Signature of quadratic form, 296, 298
- Similarity of matrices, 67
- Singularity, 143
- Smirnov, 171
- Space, coefficient, 232
  - decomposition of, 177, 248
  - dilatation of, 287
  - euclidean, 242, 245
    - extension of, to unitary space, 282
- factor, 183

- Space, rotation of, 287
  - unitary, 242, 243
  - as extension of euclidean, 282
- Spectrum, 96, 272, 273; 53
- Spur, 87
- Square(s), independent, 297
  - positive, 334
- Stability, criterion of, 221
  - domain of, 232
  - of motion, 125
  - of solution of linear system, 129
- States, essential, 92
  - limiting, 92
  - non-essential, 92
- Stieltjes, theorem of, 232
- Stodol, 173
- Sturm, theorem of, 175
- Sturm chain, 175
  - generalized, 176
- Subdiagonal, 13
- Subspace, characteristic, 71
  - coordinate, 51
  - cyclic, 185
  - generated by vector, 185
  - invariant, 178
  - vector, 63
- Substitution, integral, 143, 169
- Suleimanova, 87
- Superdiagonal, 13
- Sylvester, identity of, 32, 33
  - inequality of, 66
- Systems of differential equations, applica-
  - tion of matrices to, 116ff.
  - equivalent, 118
  - reducible, 118
  - regular, 121, 168
  - singularity of, 143
  - stability of solution, 129
- Systems of vectors, bi-orthogonal, 267
  - orthonormal, 245
- TRACE, 87
- Transformation, linear, 3
  - of coordinates, 59
  - orthogonal, 242, 263
  - unitary, 242, 263
  - written as matrix equation, 7
- Lyapunov, 117
- Transforming matrix, 35, 60
- Transpose, 19, 280
- Transposition, 18
- UNIT SUM OF SQUARES, 314
- Unit sphere, 315
- Unit vector, 244
- VALUE(S), characteristic, maximal, 53
  - extremal properties of, 317
  - latent, 69
  - left and right, of function, 81
  - proper, 69
- Vector(s), 51
  - angle between, 242
  - bundle of, 183
  - Jordan chain of, 202
  - complex, 282
  - congruence of, 181
  - extremal, 55
  - inner product of, 243
  - Jordan chain of, 201
  - latent, 69
  - length of, 242, 243
  - linear dependence of, 51
    - test for, 251
    - modulo  $I$ , 183
  - linear independence of, 51
  - norm of, 243
  - normalized, 244; 66
  - null, 52
  - orthogonal, 244, 248
  - orthogonalization of sequence, 256
  - principal, 318, 338
  - proper, 69
  - projecting, 248
  - projection of, orthogonal, 248
  - real, 282
  - scalar product of, 242, 243
  - systems of, bi-orthogonal, 267
    - orthonormal, 245
  - unit, 244
- Vector space, 50ff., 51
  - basis of, 51
  - defect of, 64
  - dimension of, 51
  - finite-dimensional, 51
  - infinite-dimensional, 51
  - nullity of, 64
  - rank of, 64
- Vector, subspace, 63
- Volterra, 133, 145, 147
- Vyshnegradskii, 172
- WEIERSTRASS, 25

